Loading the Dice
The Art of Crafting Successful Board Games

Meg Czabala, Rohan Daryanani, Michael Lisa, J Angel Ramirez, & Emma Wilson

Introduction

The focus of this project was to analyze the factors that impact the Geek Ranking and number of owners of a board game on Board Game Geeks (BGG). Machine learning models were created to predict if a game will rank well or not in the Geek Ranking and determine the number of people that will report owning the game. These models have implications that can allow game publishing companies to have a good measure of how well a game would do prior to submitting it to the market. In this way, we can make sure that our brand puts out consistently high quality games.

As determining the success of a board game is a difficult task, given the number of variables that can play into a user's preference, this problem is well suited for a machine learning method over other simpler methods. We decided that the best way to determine success of a game, given available information, is to look at the average user ratings for a game as reported by BGG, a popular website where board game enthusiasts rate and give opinions on their play experiences. Since the average review is reported on a scale of 1 to 10, machine learning will allow us to predict a discrete number for this value through regression methods. This will give us a good place to start in determining the success of a game. The data consists of several factors about the game, most notably a list of game mechanics that the game uses. We used these factors as predictors to determine the type of games users of BGG like the best.

To further predict success, we decided to set another target performance indicator of reported ownership to determine if a game is successful or not. This allows us to derive an estimate of how many copies of the game have been sold. Additionally, this is predicted with regression based methods. Given these two factors, we can predict a degree of success that a board game can achieve.

It should be noted that this data source does present some potential bias. The average user of BGG is likely a person who is much more interested in and more critical of complicated board games. This means that this data will not necessarily represent the opinions of the average consumer, but rather board game enthusiasts. Given this it is important to note that this method is to be used with the intention of targeting board game enthusiast customers, rather than the general population.

Related Work

Board Game Analytics is not necessarily a novel concept, but not all types of analysis have been done on it before. Many of the analyses that have been done have been related to the rankings of companies that produce the games with the highest ratings. Another popular style of analysis is on the board games themselves: the analytics of the strategies within each game.

In the work *Are Boardgames Getting Better? An Empirical Analysis*, the authors Chris Wray and Jeff Lingwall test multiple hypotheses to determine what led to an increase of board game ratings in the years 1994, 2004, and 2014. Some of the tests looked at the effect of weight, theme, and rating inflation. For weight, they defined weight being correlated with how complex the game happened to be, and there was no clear evidence that this was a deciding factor for rating through the years. For themes, some were utilized more in 2014 than they were in 1994 such as horror and fantasy instead of war games, but that doesn't necessarily give individuals the motivation to rate all games much higher or much lower than normal. While they did not back this hypothesis with data, they noted that rating inflation is due to people in 2014 thinking that a B is a slightly above average grade when it 1994 that could be viewed as a fairly high ranking for the game. Overall, they mention that games may just be getting better as a whole, and this is a nod to how games have been able to evolve overtime.

In *What Makes a Board Game Good: Exploratory Data Analysis of Games on BoardGameGeek.com*, Sam Beardsley scraped ranking data for 19,019 games from the website BoardGamesGeek.com to understand what makes a board game have a good ranking. In recent years, there has been a resurgence in popularity of board games. According to *Global board games market value from 2017 to 2023* by Statista, the board game industry has been steadily growing since 2017 and it is expected for that trend to continue. Statista published this data in 2019 before the COVID-19 pandemic, even though it predicted a constant increase it did not account for the surge in popularity due to the Stay-at-Home order to help mitigate with COVID cases. Beardsley published his article in 2020 and states that tabletop games are expected to increase by 4.8 billion by 2023 and that the US card and board game market alone is predicted to increase to 5 billion by 2025.

Beardsley's main purpose with this article is to understand what characteristics make a game's ranking higher to help companies understand what characteristics they should focus on in order to develop a good game. When looking at different factors, Beardsley found that in most cases lower rating games tended to fall within the children's category. This does not mean that children's games are bad, but that typically games that are rated higher tend to have strategy and more player choices than by dice rolls or random chance.

In *An Analysis of Board Games* found on Dinesh Vatvani's Github account from 2018, we found some interesting insights that were just worth noting. They looked at similar factors like the themes and general change in ratings. The complexity piece was helpful because it showed a clear increase in complexity as time went on. Also, they depicted the most popular mechanisms used in games, and dice rolling was by far the most used mechanism.

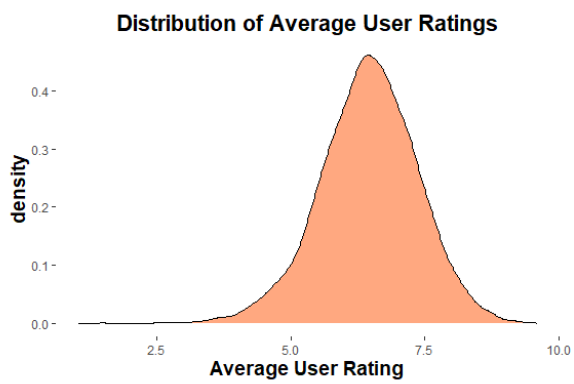
Our approach is most similar to that of Beardsley's analysis; however, instead of just focusing on the game's rating average, we decided to do a second analysis focusing on the number of users that own the games. This would be an improvement upon Beardsley's analysis as it provides a deeper understanding of what mechanics people enjoy when choosing games, and what makes people rank games higher.

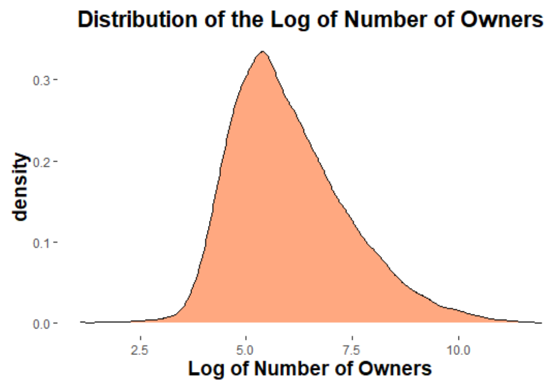
Data Description

The dataset that was used for the problem was acquired through the IEEE and shows the average user ratings and several factors for about 20,000 board games that have been rated by at least 30 people each on BGG. The data on the dataset is current as of February 2021. The columns the initial dataset includes are:

- Board game id (ID)
- Name
- Year published
- Minimum number of players recommended (Min Players)
- Maximum number of players recommended (Max Players)
- Playing time
- Recommended minimum age of players (Min Age)
- Number of users that rated the game (Users Rated)
- Average rating received by the game (Rating Average)
- BGG rank
- Average complexity value of the game (Complexity Average)
- Number of BGG registered owners of the game (Owned Users)
- Mechanics used by the game (Mechanics)
- Board game domains that the game belongs to (Domains)

Each game includes a number of mechanics that influence the experience; for example dice rolling, card drafting, flicking, hot potato, or negotiation. Games also include a number of domains that explain the type of the game, such as family games, strategy games, war games, or party games. For the pre-processing of the data, we separated the Mechanics column and Domains column into columns of dummy variables in order to encapsulate which mechanics potentially have an impact on user ratings. We created columns that are a sum of the number of mechanics and the number of domains in each game in order to have a measure of potential game complexity that is not a user evaluation. Then we removed rows that had no mechanics in them. Lastly, we standardized the names of the columns to have consistency. After cleaning the dataset, we ended up having 18,745 samples and 204 variables. For the analysis of the number of owners, we decided to use a logarithmic transformation to normalize the right-skew of this dependent variable. For both analyses, we split 80% of the dataset for the training data, and the other 20% was used as the test data. On top of that, cross validation was used for tuning the hyper parameters.





Methods

After cleaning our data, we decided to create two XGBoost models. We chose XGBoost because it has strong prediction powers, especially with large datasets. Although it is much slower than other models to train, we were ideally looking to have the greatest accuracy, and lowest errors (RMSE and MAE). It should be noted that while XGBoost does produce complex robust models and predictions, it does suffer from being difficult to interpret and visualize. Additionally, boosting can be overly sensitive to outliers which could skew our results. We also included a linear regression as an initial test of our dataset. This was used as an initial analysis to check if the data was properly prepared for a more robust method and to take a look at potential factors of importance.

Results

The initial runs of our models gave good results, but we decided these could be further improved by tuning the hyper parameters involved in the XGBoost methods using cross validation. After running cross validation to test the best combination of parameters, we determined the following values to use in our final models:

	eta	Max depth	Min child weight	gamma	subsample	colsample_bytree
Ratings Model	0.1	10	7	0.1	0.9	0.6
Owners Model	0.1	10	10	0.00	0.9	0.6

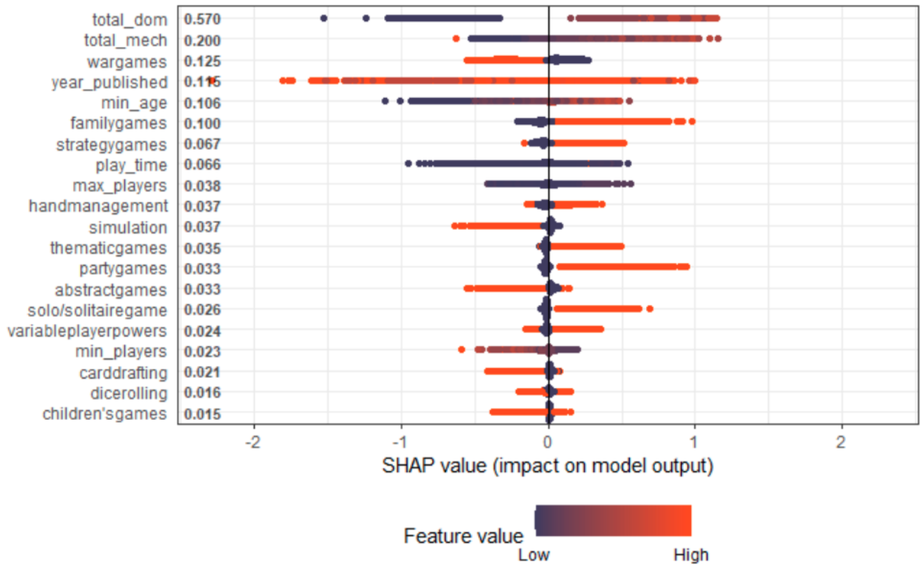
The designed models output a predicted number for the rating and ownership for a supplied set of characteristics of the board game. After this, we obtained error metrics for our model based on the test data set. Since we were performing regression analysis, our goal was to minimize RMSE and MAE when tuning the model. RMSE is the root mean squared error which measures the distance between the actual values in the test dataset and the predicted values by the model. MAE is the mean absolute error which measures the average of the absolute errors.

This indicates how far off each prediction is expected to be from the actual value. For ratings, we obtained a RMSE of 0.72 and a MAE of 0.54. For owners we obtained a RMSE of 804.9 and a MAE of 334.6. These are good prediction results, as for ratings the model is off by only 0.54 on average (on a scale of 0 to 10) and for owners the model is off by only 334.6 owners on average (for a value that can have several thousand reported owners). While these final models are relative successes, we did face challenges. Runtime was an anticipated problem that came to be a reality; however, by utilizing parallel processing and multiple computers, we were able to cut down the runtime exponentially. Also, the likelihood of having found the optimal models is low despite the low error metrics. We could continue to tune the model by adding more values to each cross validation to minutely improve the error. That being said, we used a large range of values and ensured that the optimal value was well within the set and not an edge value. While not perfect, the XGBoost models contain information that can be applied to the board game industry.

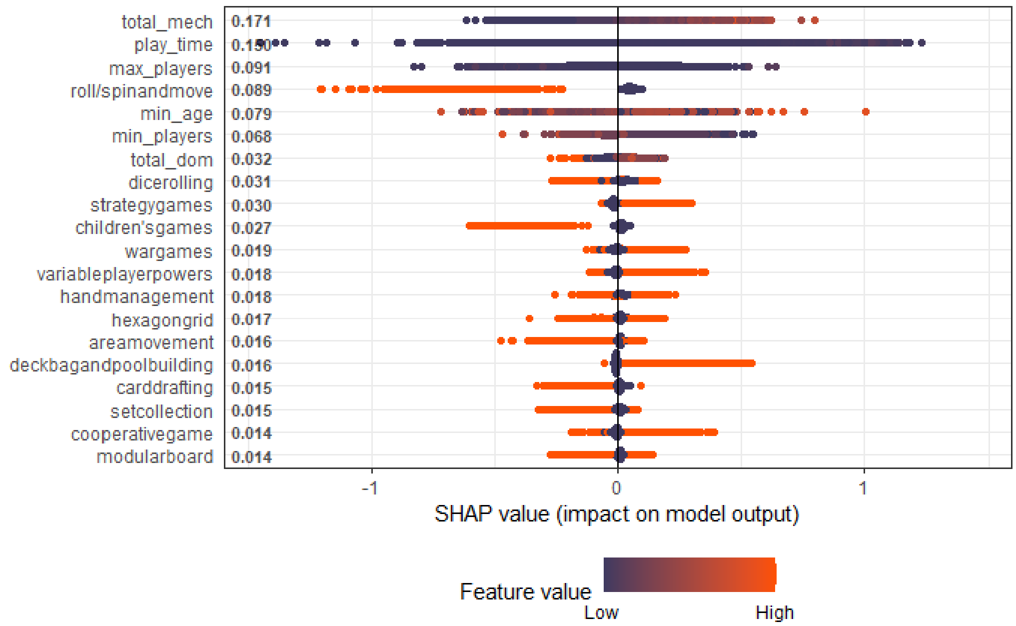
Discussion

After performing the XGBoost Analysis, we developed some insights that could help board game creators. Our recommendations depend on whether the board game creators are more interested in selling the most amount of games to have the highest amount of revenue, or if the board game creators want a highly rated game with a loyal following.

If the board game creator wants to sell the most amount of games to maximize profit, they should develop a simple, thematic, family-based party game, where parents can play with their children (e.g. Game of Life). Some reasons for this might be because the games are simpler to play, and thus there is a low knowledge barrier to entry to start playing the game. Usually less complicated games have less mechanics, and thus, on average, are less expensive to make. This allows creators to sell them at lower prices and so people are more inclined to purchase the game.



On the other hand, if the board game creators want to have highly rated games, they should develop a more complex game with multiple mechanisms, avoid randomness/chance (e.g. dice rolling), and cater to teenagers/young adults. Some reasons for this might be because more experienced players would rather the game depend on their skill, rather than through chance or luck. As players spend more time playing board games, they start wanting to play more complex games to challenge themselves. On average, the majority of board game enthusiasts are teenagers and young adults, and because of this, they desire to play games that are more appropriate for their age.



Conclusion and Future Work

In conclusion, our models of predicting board game ratings and reported ownership on BBG produce an accuracy that can be useful for board game companies to predict if a game they wish to be published will be well received and successful. We further found that factors to focus on to receive high ratings include a large number of mechanics, reduction of elements of chance, and a focus on strategic elements. Factors that contribute to higher reported ownership are accessibility to families and parties, avoiding complicated themes such as wargames and simulation, and the ability to play the game with one player. These design elements can be used by board game companies to create better games for the dedicated fans that use BBG.

Other considerations to keep in mind would be to include the MSRP of board games as another variable in our models. By adding this additional variable it will allow us to categorize by different price ranges and to compare board games in different price ranges. Another consideration could be to use and analyze a less skewed dataset. Our current data comes from a sort-of-niche website that consists of board game enthusiasts ratings. Adding the extra layer of regular customer ratings will allow us to be closer to a normal distribution allowing us to perform more accurate models and at the same time reducing bias. One example would be to add different reviews from either Amazon or Walmart. This will give us a less biased dataset as well as allowing us to extract the MSRP from the source itself.

Bibliography

Beardsley, S. (2020, August 23). *What makes a board game good: Exploratory Data Analysis of games on boardgamegeek.com*. LinkedIn. Retrieved November 17, 2021, from https://www.linkedin.com/pulse/what-makes-board-game-good-exploratory-data-analysis-games-beardsley/.

Bedford, E. (2019, August 9). *Global Board Games Market Value 2017-2023*. Statista. Retrieved November 17, 2021, from https://www.statista.com/statistics/829285/global-board-games-market-value/.

BoardGameGeek. (n.d.). Retrieved November 21, 2021, from https://boardgamegeek.com/

Samarasinghe, D. (2021, July 6). *BoardGameGeek dataset on board games*. IEEE DataPort. Retrieved November 21, 2021, from https://ieee-dataport.org/open-access/boardgamegeek-dataset-board-games.

Vatvani, D. (2018, December 8). *An analysis of board games: Part I - introduction and general trends*. Github. Retrieved November 21, 2021, from https://dvatvani.github.io/BGG-Analysis-Part-1.html.

Weisholtz, D. (2020, December 5). *How classic board games are bringing families closer during the pandemic*. Today.com. Retrieved November 21, 2021, from https://www.today.com/popculture/board-games-enjoy-surge-popularity-during-pandemic-t202377.

Wray, C., & Lingwall, J. (2017, March 22). *Are Boardgames Getting Better? An Empirical Analysis*. The Opinionated Gamers. Retrieved November 17, 2021, from https://opinionatedgamers.com/2015/08/17/are-boardgames-getting-better-an-empirical-analysis/