Abby Ananthakrishnan
Helena Shi
Joseph Chin
Jeffrey Fulkerson

How to Break Into the Board Game Market

Description of Project Goals

Our Dataset

The Board Game dataset was found on Kaggle.com, but the data was originally collected from the website https://boardgamegeek.com/ via API. We are running analysis on a dataset of 5000 different board games. It includes the following information we determined is relevant to our analyses on every board game:

- Name
- Minimum required players
- Maximum players allowed
- Average time per play session
- Average rating (from voters)
- "Geek" rating (special rating variable from boardgamegeek.com)
- Number of votes per game from the survey
- Minimum age recommended
- Mechanics (strategies employed to win)
- How many voters already own this game
- Category (theme and setting)
- Designer
- Weight (a 0-5 numerical measure of game complexity)

We are trying to determine what sorts of board games are most likely to be successful in this highly competitive market. Specifically, we want to answer the following two questions:

1. If a business owner is trying to create a new board game, what features can we recommend them to produce the highest ratings and popularity?
2. Using cluster analysis, if a user inputs certain characteristics of their favorite game, can we recommend a similar game for them to try?

Why Board Games?

According to Statista.com, "In 2016, the global board games market was valued at approximately 3.2 billion U.S. dollars." The forecast in 3 years, for 2021, shows a projected value of 8.12 billion dollars, which highly indicates that the board game industry is growing at an exponential rate.

By combining the nostalgia of family tabletop evenings with innovative advances in design, games like *Ticket to Ride* have huge followings, and board game lovers have created their own unique culture. Designing a game people want to play represents a significant challenge because finding an objective metric of comparison is non-trivial, but successfully publishing a popular board game can represent a significant return on investment, and inclusion in a lucrative community of game enthusiasts.

Helping designers publish more entertaining board games also directly benefits the public. The study of games is a study on maximizing engagement and joy, for individuals and for friend groups. Board games are common shared activities between groups of people, so we only stand to benefit from learning how to make them better.

Exploratory Analysis

Weight/Complexity

One of the most difficult variables to quantify is complexity, referred to as "weight" in the dataset. It is a measure of how difficult voters found it to learn the rules and remember them during play. Unintuitively, our analysis suggests that people prefer more complex games, rather than games of low to moderate difficulty. We suspected that high complexity might create a barrier for entry that would lower overall ratings, but our model suggests it only engages them further.

Simple analysis of the top ranked games seems to confirm this, although perhaps complexity should be broken down into more specific information. Currently, it is only a numerical measure from 0 to 5. Further exploration into types of complexity (i.e. mechanics) is likely to be insightful.

Number of Players

We are very interested in how the social dynamics of particular games affects how much people enjoy them, all other things being equal. Is it more fun to play with an intimate group or a large one? According to our analysis, two players is the best model, having the option to play alone or with 2 to 4 players is better than playing alone, but playing alone is better than playing in a large group.

This is a somewhat surprising insight, because it means that games which have the option to be played solo or in groups beat out games that require two or more people. Perhaps this is because the voters are fond of games they get more use out of.

Recommended Minimum Age

The analysis of recommended minimum age seems useful. It shows clear advantage to targeting ages 13+. This makes sense, because teenagers and adults can comprehend more complex games, and at 17+ more mature content opens up to add appeal. Games like *Cards Against Humanity* would be a mature game in this category, and illustrate that there is a combination of games that are complex and games that are simple but "adult-oriented" However, we determined that this variable is likely highly correlated with the weight variable. Since it is also likely indirectly measuring two different factors (complexity early on and adult-appeal later on) we decided not to include it in our model.

Play Session Length

Part of the social dynamics of the games is how long we spend playing them. Some board games are quick, between 15 and 30 minutes. Others take hours to complete in an average play session. We even have one or two outliers that literally take days to resolve. Since play length has obvious effects on when a game will be played and whether it will be reach resolution, we view it as important to our analysis.

Board games were more popular the more time that was required to play. This suggests that most people are fonder of games that keep them engaged longer. We speculate that these games leave a more lasting impression due to their required commitments. Familiarity breeds fondness.

War Games

Game categories represent the dressings of the game, like its visuals, story, and tone. We thought these would be considerably impactful on the ratings of the game based on their role-playing and narrative appeals. The unique categories cover a wide range, but one commonality among the categories of the top rated games is war. World War I, World War II, the Civil War, Korean War, and others were very popular.

In future analyses, we would like to come up with an informed bucket system to group these categories into smaller ones, like War. We think this could yield some valuable insight into appealing to fans of particular thematic games.

Card Drafting

Mechanics was the column representing rules and strategies employed during play. These are important qualities of the board game because they make up the core gameplay, and in theory the most significant appeal of the game. It is a difficult category to analyze, however, because of two considerations: 1) a truly accurate analysis will measure the combinations of mechanics, which will be much more computationally intensive than our independent approach, and 2) there is not data available on how well

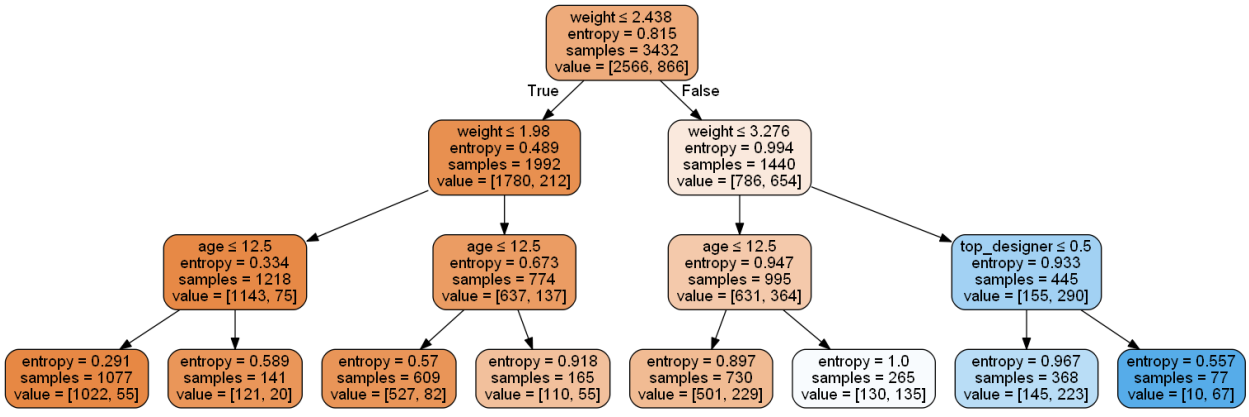
implemented a particular board game's mechanics are compared to another's.
However, from our independant analysis we observed that card drafting games are on the rise. Card drafting is a mechanic in which there are shared pools of in-game resources and players make tactical decisions about which to take for themselves, either to advance their own position or to deny the resource to another player.
In future analyses, we would like to delve more into the interdependent nature of game mechanics.

Solution and Insights

We used a decision tree (entropy) after studying our univariate analyses. After a few attempts to tune its depth, we arrived at a tree of depth three for general use. It says that the most important factors in making a popular game are:

- **High Weight** : Make your games challenging!
- **High Minimum Age Requirement** : Aim for ages 13 and up, including mature games

This result surprises us somewhat, because we suspected that mechanics or category would play more significant roles. However, mechanics is likely correlated with complexity to some degree (some strategies are more difficult to implement than others) so it is likely still involved to a degree.
In the future, we would like to run a deeper analysis, and create a deeper tree, to determine if we can get a higher accuracy by allowing more variable consideration. With our current model, we have an overall accuracy of 77% and a 26% of correctly guessing what kind of game will be popular.



For more details, go to:
https://github.com/helenashi95/Deep-Project-Group-12