Aaron Huang
Springboard Capstone 1 Project Proposal

Title: Board Game Recommender

Background:

In the past few years, the board gaming hobby has seen a large and steady growth. Retail sales of hobby board games in the USA alone grew 13% from $305 million in 2016 to $345 million in 2017, and the global board game market is forecasted to reach $12 billion by 2023. This growth is particularly evident on crowdfunding platforms: in 2018, board game projects accounted for almost a third of the $608 million raised by all projects on Kickstarter. However, this rapid growth has introduced confusion among distributors, retailers, and consumers. With thousands of new games being released each year, distributors do not know which games to order and retailers do not know which games to stock.

The goal of this project is to develop a model to accurately predict what board games an individual would enjoy based on the games that others with similar tastes enjoy.

Client Interest:

The client for the project is the board game retailer. Many retailers stock hundreds or thousands of board games, and being able to accurately recommend products to a shopper would improve the shopper's experience and increase sales. Additionally, a retailer could use a recommender to determine additional games to stock based on which games in their current stock are good sellers.

Data Source:

The data source for this problem is BoardGameGeek.com. This website contains a database of over 100,000 games with detailed information and user ratings for each game. Additionally, the website is highly active, with over 2 million registered users and attracting over 4 million unique visitors per month. Their database can also be queried by a public API, allowing for relatively straightforward gathering of game information and ratings.

Approach:

To solve this problem, a variety of item-item and user-item collaborative filtering methods will be employed beginning with simpler approaches such as prediction using cosine similarity or Pearson correlation and moving on to more detailed approaches involving clustering and matrix factorization models.

Deliverables:

The deliverables from this project will include a project code and a written analysis of the process and outcomes.