

Predicting Board Game Ratings

Bella Ball

Research Question

Research Question

Can a board game's rating be predicted by its features, such as play time, number of players, category, mechanics, and the release year.

Context

The website BoardGameGeek tracks information on board games and users' opinions on games. Using this information could help publishers choose the best game to promote.

Hypothesis

There is a significant relationship between a game's features and rating, making it possible to predict a game's rating based on its features.

Data Analysis Process

Data Collection

Data Cleaning

Summary Statistics

Regression Model

Optimize Model

Data Collection

BoardGameGeek Reviews

Data Card Code (51) Discussion (5) Suggestions (0)

▲ 142

Code

Download



games_detailed_info2025.csv (110.06 MB)

Download



Detail Compact Column

10 of 52 columns

About this file

Suggest Edits

This file does not have a description yet.

#	type	id	thumbnail	image
0	27.8k	1 unique value	27762 unique values	27759 unique values
0	boardgame	13	https://cf.geekdo-images.com/PyUo19QxBnZQCJqZI6bmSA__thumb/img/virV2Bm82Dq17gh-LZScBwqByik=/fit-in/2...	https://cf.geekdo-images.com/PyUo19QxBnZQCJqZI6bmSA__original/img/g11AF48C6pLizxWPAqdUEeKltQ=/0x0/f...

Data Explorer

Version 4 (5.41 GB)

- 2020-08-19.csv
- 2022-01-08.csv
- bgg-15m-reviews.csv
- bgg-19m-reviews.csv
- bgg-26m-reviews.csv
- games_detailed_info.csv
- games_detailed_info2025.csv

Summary

7 files

146 columns

Data Cleaning

Import Data

Check For Null/Duplicate

Drop Columns

Convert Columns to Counts

One-Hot Encoding

Combine Dataframes

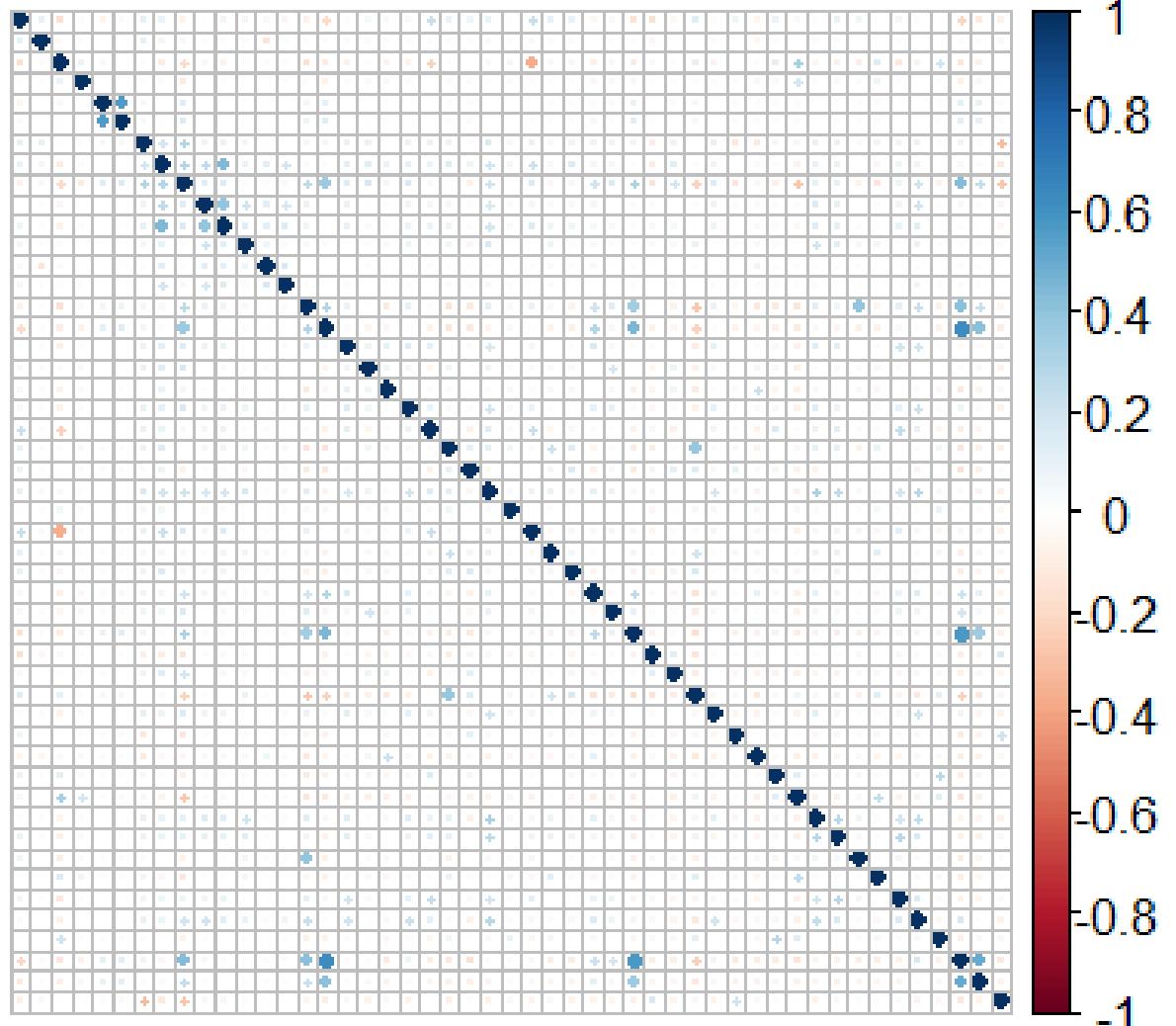


Summary Statistics

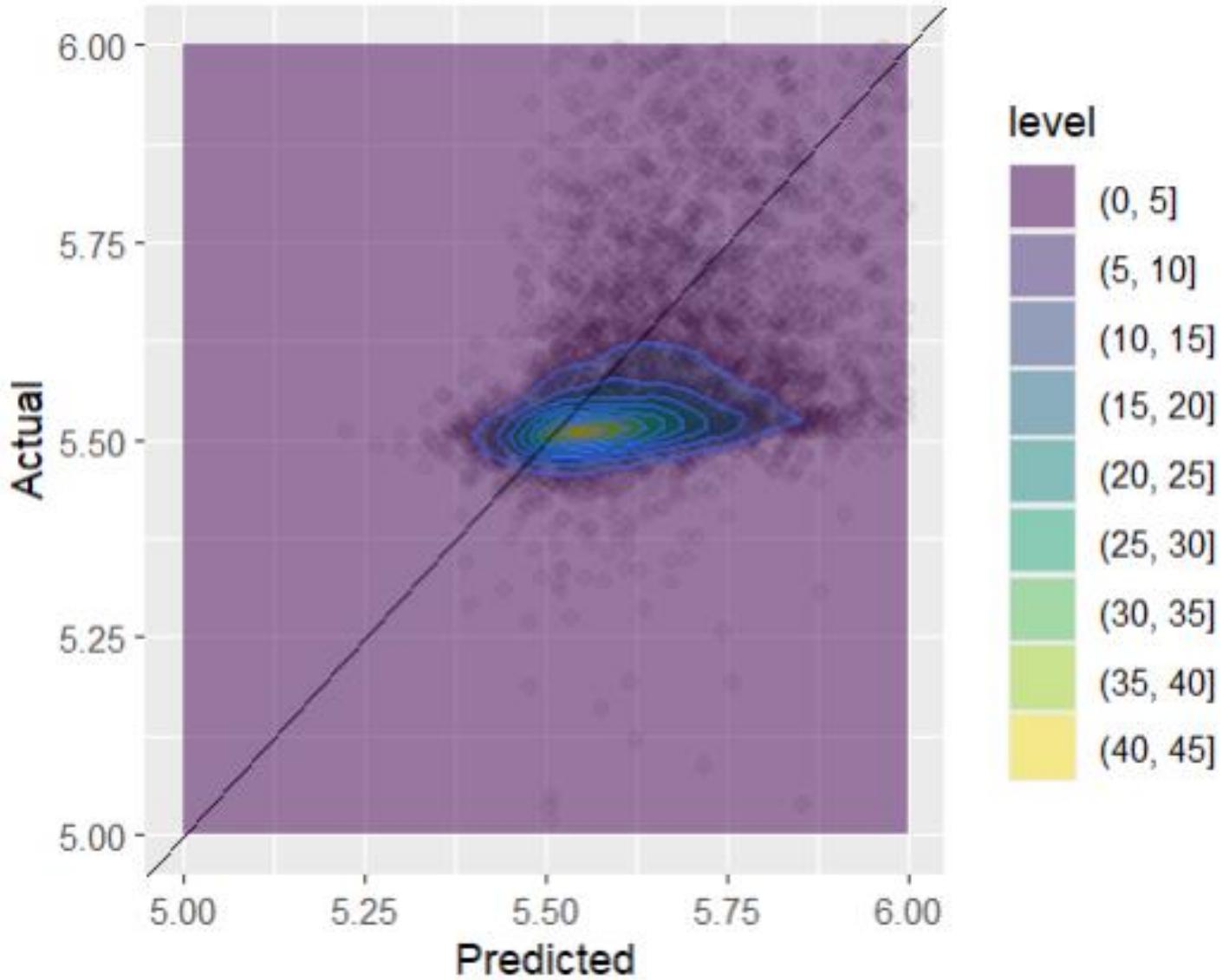
Variables	Minimum	Maximum	Mean	Mode	Standard Deviation
yearpublished	-3500	2026	1989.18	2014	205
minplayers	0	10	1.96	2	0.71
maxplayers	0	999	5.94	4	13.73
minplaytime	0	60000	56.94	30	380.14
maxplaytime	0	84000	91.94	45	696.45
minage	0	21	9.68	10	3.65
bayesaverage	3.66	8.41	5.67	5.54	0.35
expansion_count	0	570	1.44	0	8.13
accessory_count	0	78	0.36	0	2.10
integration_count	0	75	0.31	0	1.86
compilation_count	0	15	0.06	0	0.34

Regression Analysis

- Adjusted R²
 - .3671
- P-value
 - <0.0001



Optimizing the Model



Final Regression Formula

Bayes_average = 5.29 + 0.05(accessory_count) + 0.09(averageweight)
- 0.11(Wargame) + 0.05(implementation_count) + 0.15(Solo/Solitaire
Game) + 0.1(Open Drafting) + 0.09(Area Majority / Influence) +
0.09(compilation count) + 0.06(Hand Manangement) +
<0.01(expansion_count) + <0.01(minage) + 0.06(Set Collection) -
0.05(Card Game) - 0.08(Roll / Spin and Move) + 0.07(Economic) +
<0.01(integration_count) + 0.06(Deduction) + <0.01(yearpublished) +
0.04(Variable Player Powers) + 0.05(Tile Placement) +
0.05(Simultaneous Action Selection) + 0.04(Animals) +
0.04(Cooperative Game) - 0.04(Abstract Strategy) - 0.04(Humor) +
0.02(Party Game) - 0.03(Adventure) + 0.02(Modular Board) -
0.03(Miniatures) - 0.02(Children's Game) + 0.02(Action Points) +
0.02(Area Movement) - 0.02(World War II) - 0.02(Fighting) +
0.02(Grid Movement)

Findings

MSE

•0.0805

Adjusted R²

•0.3671

P-value

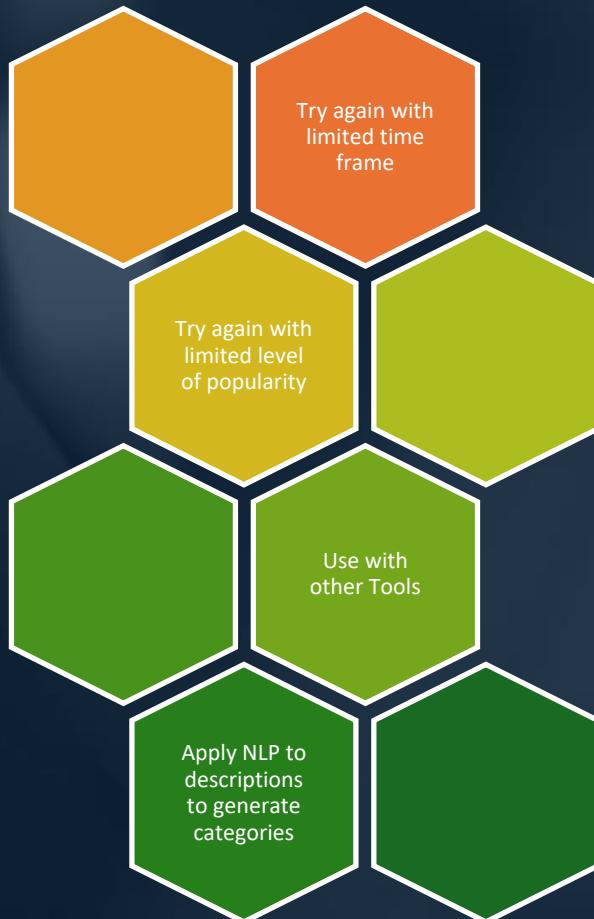
•<0.0001

Limitations

Limited
accuracy

Complex
model

Further Analysis



Recommended Actions



- Solo
- Complex
- Area majority
- Hand management
- Set collection
- Deduction

- Wargames
- Roll / spin and move
- Card games

Expected Benefits



INSIGHT INTO CUSTOMER
PREFERENCES



DATA-DRIVEN DECISIONS



INTENTIONAL
PROMOTIONAL DECISIONS



IMPROVED MODELS

References

Board Game Geek. (n.d.). *Board Game Geek*. Retrieved from ratings:
<https://boardgamegeek.com/wiki/page/ratings>

Elteren, J. V. (2025, January). *Board Game Geek Reviews*. Retrieved from Kaggle:
https://www.kaggle.com/datasets/jvanelteren/boardgamegeek-reviews?select=games_detailed_info2025.csv