Data Analysis of Board Games' Fashion Trends and Relationships between Their Successes and Attributes

Zhongxing Du

Dalian University of Technology, Dalian, 116081, China

Abstract. From the tic-tac-toe, an ancient chess game, which was invented thousand years ago, to modern wargames that take weeks to finish, board games have long been defined as a tool to create a platform for entertainment, education, and training of strategy in fields such as military affairs and economics. Development of technology is worthy of a rapier to the field of board games, both bringing new-published board games to tables worldwide and replacing them with more convenient entertainments such as video games or computer games. Towards this end, this thesis presents an analysis-driven review of board game domains and play-related attributes based on data from the world's largest board game website - BGG. The analysis expects to make a clear view of the changes in users' favors by comparing recent and former board games. And further investigate the relationships between their successes and inner attributes.

Keywords: Board Games; BGG; Data Analysis; Prevalence.

1. Introduction

It is believed that the very first board game 'Senet' was invented before 3000 B.C, being buried in ancient Egyptian ruins [1]. Then after 5000 years of development, not less than tens of thousands of board games were designed, published, and spread over civilizations and histories.

After thousand years of revolution, 8 domains and more than 180 mechanics have been created and many remain popular all over the world. Since the beginning of the 21st, with systematic design methods and the assistance of advanced tools, board game publishment received rapid growth.

Under such circumstances, studying the changing process of board games as well as their inner relationships between popularity and attributes will be helpful for board game publishers and practitioners to grasp players' preferences and improve mechanics in their new designs, which is meaningful for predicting future trends in board game fields.

In this thesis, by analyzing a dataset that collects board game data from BGG, the largest online board game forum in the world, the fashion trends of board games and relationships of their success in the past decade are revealed, with results and discussions that can help to comprehend the changes and trends of more than 20000 extant board games. Also leaving potential and possible direction for further investigations and trials.

The paper is organized as follows. In section 2, we introduce related works on board games. A description of the data set used and a data dictionary is presented in section 3. Section 4 presents detailed analysis, results, and discussions based on aforementioned parts. Finally, an overall conclusion is drawn in section 5.

2. Related Work

There have been many studies focusing on board games's effects on fields beyond entertainment, like preschooler education [2], linguistics [3] and medical science [4], yet few researchers have explored board games' themselves attributes and their popularity. A review of board game mechanics interactions presented in 2021 draws conclusions on connections of gaming mechanics and some gaming experience qualities by analyzing top 10000 games on BoardGamesGeek dataset, showing the current most popular mechanics in 2021 to readers. However, their exploration of relationships used static rankings which lacks a dynamic description of board games' popularity trends, especially the changes of popular domains, mechanics over recent years.

Meanwhile, some researchers have explored particular attributes of board games. Samarasinghe D, Barlow M, et al. have discovered interactions of their mechanics with data driven methodology [5]. However, most researchers focus on attributes of a larger range of games rather than typical type of board games, like Dietrich T and his team who analyzed gaming attribute preference in marketing programmes [6].

Inspired by previous works, this paper dynamically analyzes fashion trends of board games by comparing different time periods of board game history, as well as discovers the relationships between their success and attributes. This paper not only integrates effects of attributes on popularity, but also presents trends of outmoded and accelerating board games' designs, to better assist people witness the variability of board game field throughout the years.

3. Data Set

A board games dataset on 'BoardGamesGeek' is used in this thesis, which provides information on extant board games from 3000 B.C. to 2021H1, and was updated online. A total number of 20343 board games with the most valuable information were collected, despite there being more than 100,000 games available on the BGG website which were lack of information that may disturb the research. A review of interactions of board game mechanics was presented in 2021, concluding connections between board games mechanics with some gaming experience A minute description of attributes in the database that are involved in the research is demonstrate in Table 1, also to make it feasible for readers to reference current information on board games.

Table 1. Data Dictionary

	Attribute	Explanation
1	ID	Unique BoardGamesGeek ID [sorted in descending order according to comprehensive ranking on BGG]
2	Name	Name of the board game
3	Year Published	Year published [range: 3500 B.C. – 2021 A.D.]
4	Min Players	Minimum suggested players
5	Max Players	Maximum suggested players
6	Play Time	Average play time suggested by game creators [Numeric
7	Min Age	Age rating
8	Users Rated	Amount of BGG users who reviewed the game
9	Rating Average	Average BGG player rating
10	BGG Rank	BoardGamesGeek ranking
11	Complexity Average	Average BGG community complexity rating [Numeric 1-5,0: uncategorized]
12	Owned Users	Amount of BGG users who said they own the game
13	Mechanics	BGG Game Mechanics (182 separate)
14	Domains	BGG community voted game subgenres the BGG site

Even the dataset has included over 20000 relative renowned board games, there are still duplicated or missing data in the dataset, thus it is essential to clean the data and choose the appropriate lines to analyze.

The following Figure 1 illustrates the distribution of the average rating and BGG rank of the original data set of 20,343 board games totally.

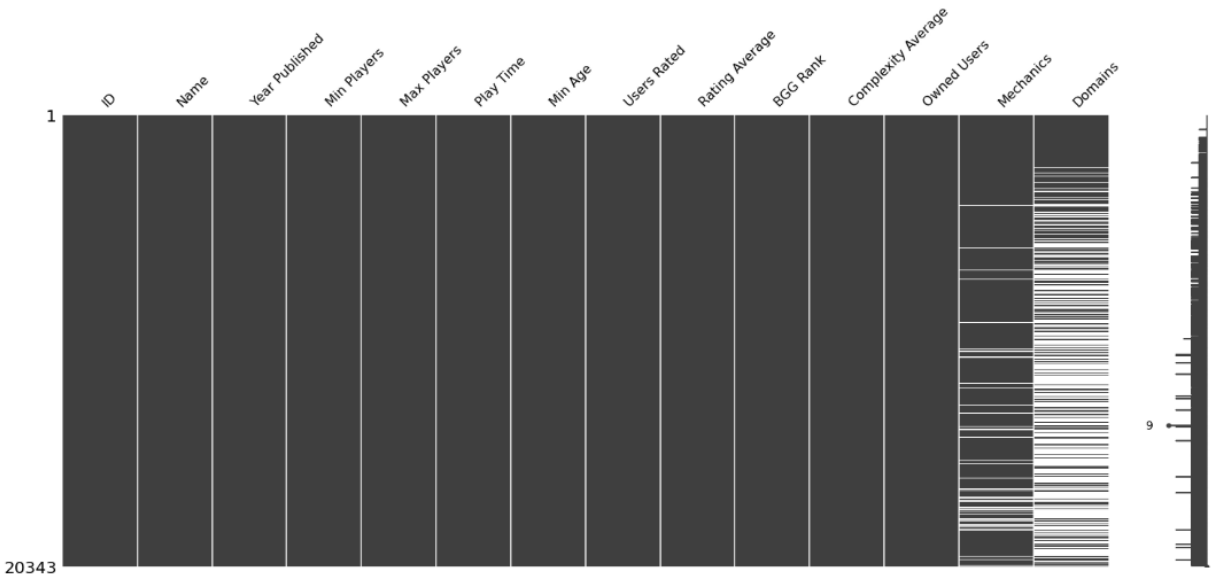


Figure 1. Distribution of average scores and BGG rankings of 20343 board game raw data sets

Table 2. Available Values

Column	Non-Null	Type
ID	20327	Float64
Name	20343	Object
Year Published	20342	Float64
Min Players	20343	Int64
Max Players	20343	Int64
Play Time	20343	Int64
Min Age	20343	Int64
Users Rated	20343	Int64
Rating Average	20343	Object
BGG Rank	20343	Int64
Complexity Average	20343	Object
Owned Users	20320	Float64
Mechanics	18745	Object
Domains	10184	Object

In Table 2, a slight number of data are missing in *ID*, *Year Published*, and *Owned users*, while a few data are missing in mechanics. Nearly half of the board games are not divided into any domains.

Figure 2 partly explains that the null values were mainly distributed in recent years.

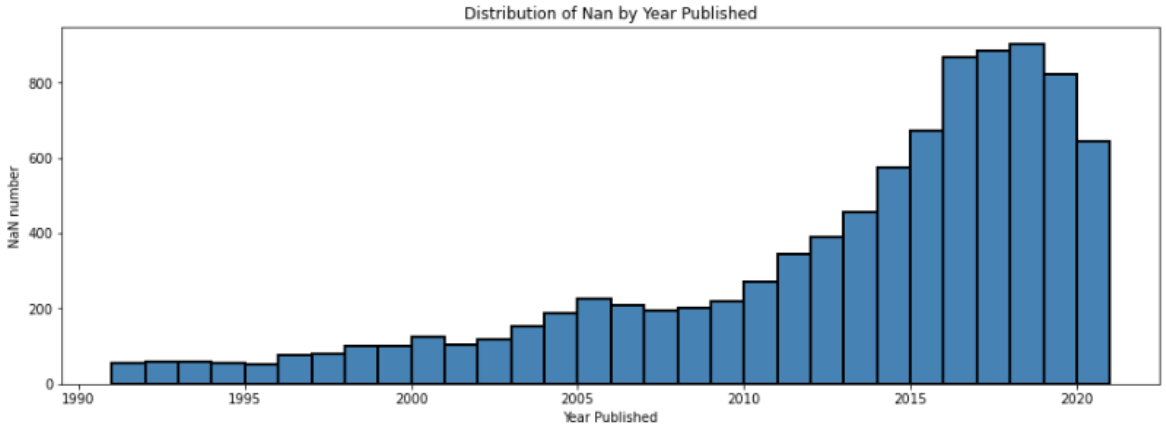


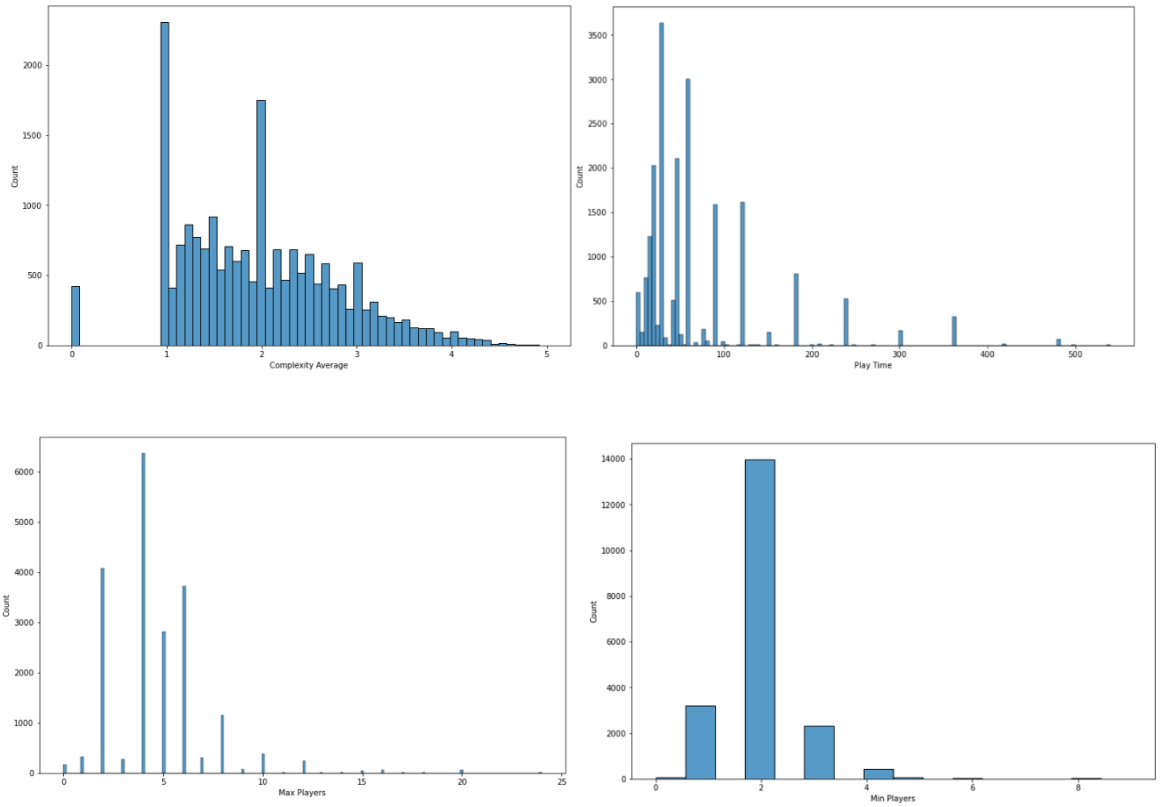
Figure 2. Null value of main distribution in recent years

This paper uses their mean values to replace the null values in Year Published and Owned Users, and empty columns in Mechanics and Domains will be labeled 'Uncategorized'. Besides, data with extreme values (Far from average for example) will normally be erased before analysis.

4. Data Analysis

A few questions were listed below for further analysis of the dataset, while the results presented are detailed evaluations mainly focused on showing: (1) Overall attributes of board games;(2) trends of modern board games designed/published in recent years;(3) Relationships between the successes of board games (measured by favors of users) and their attributes, as well as discussion and conclusion of the results.

4.1 Overall Attributes of Board Games



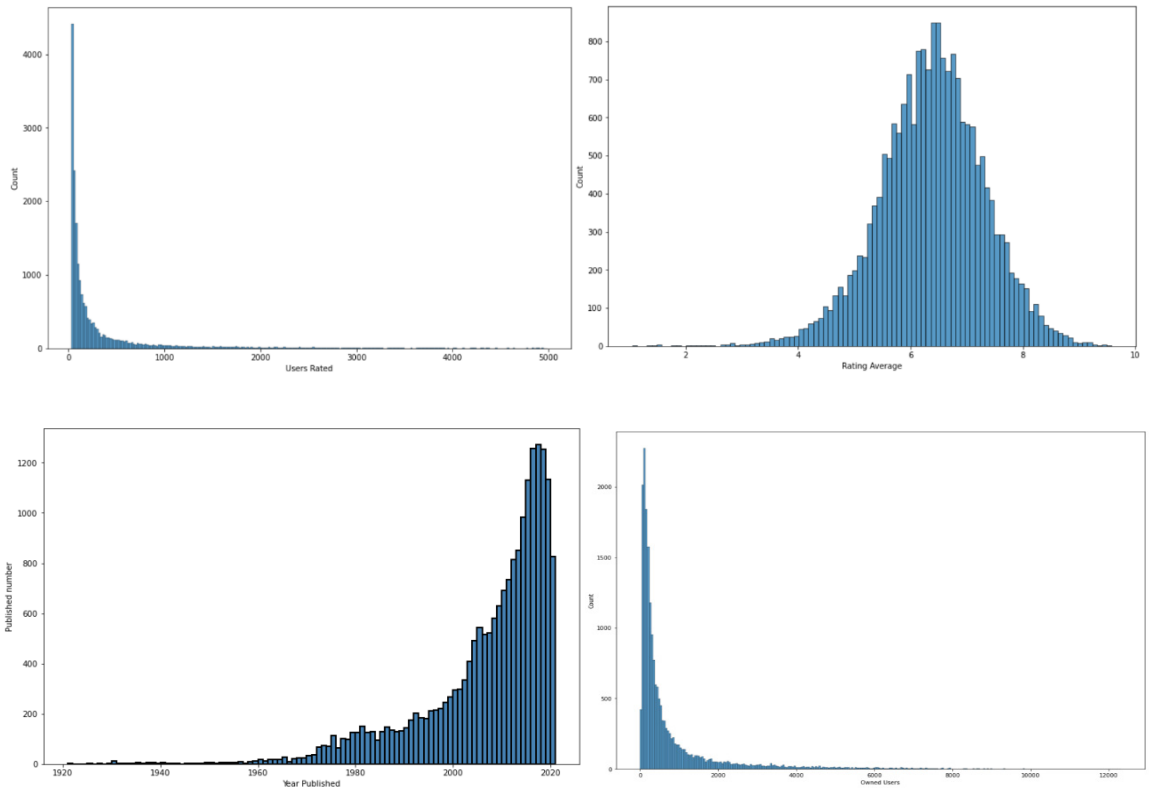


Figure 3. Total number of board games released in history

Figures 3 to Figure 10 illustrate the total number of board game released over history on the based on their basic attributes. Some extreme data are erased in the statistics respectively.

Complexity Average, *Play* Time, *Max Players* (0-25), *Min Players* (0-10), *Users Rated* (0-5000), *Rating Average* (0-10), *Year Published* (after 1800 A.D) and *Owned Users* (0-12000):

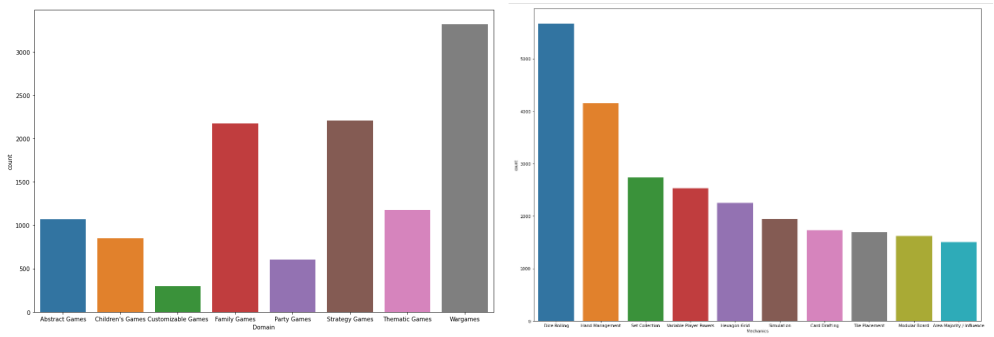


Figure 4. Individual Domains (ignoring uncategorized)

Table 3. Evaluation Table

	Year	Min	Max	Play	Min	Users	Rating	Complexity	Owned
Count	20342	20343	20343	20343	20343	20343	20343	20343	20320
Mean	1984.25	2.02	5.67	91.29	9.60	840.97	6.40	1.99	1408.46
Std	214.00	0.69	15.23	545.45	3.65	3511.56	0.94	0.85	5040.18
Min	-3500.00	0.00	0.00	0.00	0.00	30.00	1.05	0.00	0.00
Mid	2011.00	2.00	4.00	30.00	8.00	55.00	5.82	1.33	146.00
Max	2022.00	10.00	999.00	60000.00	25.00	102214.00	9.58	5.00	155312.00

According to figure 3 to figure 10, among the 20343 board games included, most of the games have complexity levels of 1-2, min players of 1-3, and max players of 2-6. Most board games have less than 1000 users rated averaging mainly 4-6 out of 10 and most of them are owned by less than 3000 users. Nearly all the games included were published after 1800 A.D, but a dramatic increase trend of board games published can be witnessed, beginning from 1960 and met its maximum in 2018, and may show a downtrend after then. Besides, the median *Year Published* is 2011, thus the year 2011 will be regarded as the boundary between previous board games and modern ones in further analysis.

Obviously, among all the existing board games, war games were published the most based on Figure11, followed by strategy games and family games of more than 2000 existing in the board game field. However, with 182 mechanics being listed on BGG, 'dice rolling' and 'hand crafting' board games have significant gaps between the others, which includes more than 5000 and 4000 respectively.

In terms of figure 2 which shows the distribution of Null numbers, nearly two-thirds of new-published board games every year are not well-categorized, i.e., two-thirds of the board games are sifted out after their publishment at any moment in history.

4.2 Trends of Modern Board Games Designed/Published in Recent Years

The historical trend of board games has been concluded in the last section, but it is often the recent market that interests board game designers, publishers as well as players. Therefore, Figure 5 show the attributes and trends of board games that were published during 2011 and 2021, which is also the period when board game markets thrived in history.

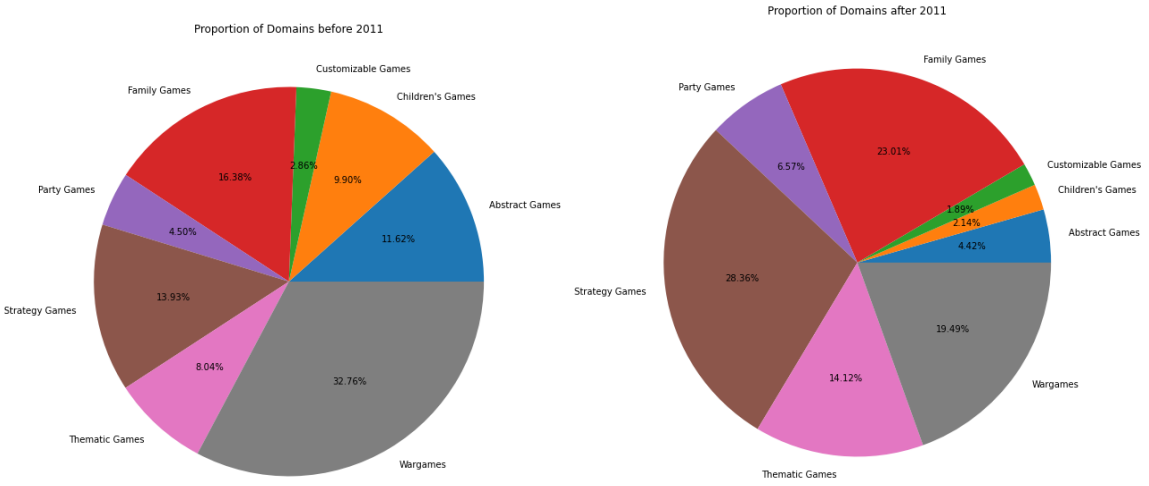


Figure 5. Properties and trends of board games released in 2011 and 2021

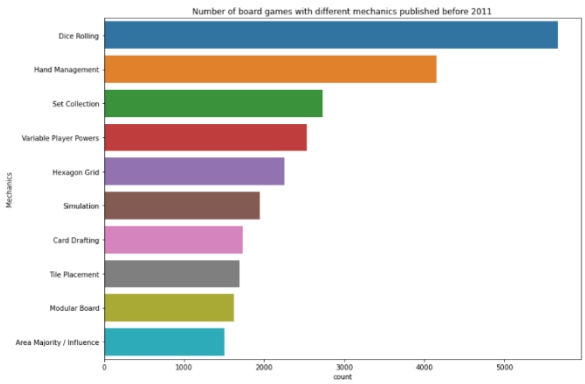


Figure 6. Number of board games with different mechanisms published before 2011

On the basis of Figure 5, significant changes are obvious to be found in the three most popular domains – Strategy Games, Family Games, and Wargames. Strategy games were published the most followed by family games. Wargames, which were the overall king of board games fell from the altar, but still had a steady publish number of more than 800. Especially, 'abstract' games (e.g., chess) and 'children' games took much less proportion of the total publishments compared to the entire history. It tends to show a trend that board game players have an older age group and prefer games with more particular and vivid playing rules compared to the past, which leads to a noticeable decrease in these two domains.

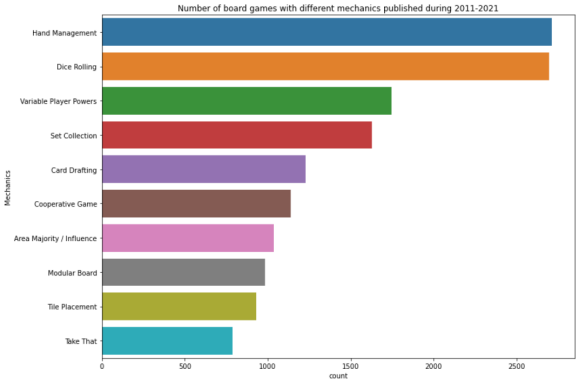


Figure 7. Number of board games with different mechanics published during 2011-2021

Similar great changes can also be found in mechanics according to Figure 6 and 7, compared to the entire board game history, most top 10 mechanics remained popular over thousands of years. Yet 'Cooperative Game' and 'Take that' are two rising stars, replacing the seats of 'simulation' and Hexagon Grid' games, which were once top 5 and 6 among hundreds of mechanics.

Table 4 illustrates top ten mechanics with the fastest-growing ratings in 2011-2021 compared with board games published before 2011, in which D-value is the rating difference value and P-increase describes the percentage of increase compared with ratings before 2011.

Table 4. The Fastest Growing Mechanics on Ratings

Mechanics	2011-2021 Rating	-2011 Rating	D-Value	P-increase
Singing	7.79	5.68	2.11	37.15%
Finale Ending	7.30	5.49	1.81	32.97%
Static Capture	7.40	5.67	1.73	30.51%
Line Drawing	7.24	5.75	1.48	25.72%
Pattern Movement	7.39	6.00	1.39	23.16%
Acting	6.99	5.68	1.31	23.06%
Mancala	7.77	6.50	1.27	19.63%
Score-and-Reset Game	8.43	7.18	1.25	17.41%
Resource to Move	7.87	6.65	1.22	18.26%
Storytelling	7.55	6.42	1.13	17.60%

These mechanics, with rapid-growing praises, show their potential and popularity in recent years.

Also, 9 new mechanics were created after 2011 which are presented in Table 5, and some of them such as 'Passed Action Token' and 'Legacy Game', have won good reputation within the period of 10 years. Furthermore, among all the modern mechanics, a decrease in rating is only witnessed in 15 mechanics, meaning progress in board game designs these years.

Table 5. New Mechanics after 2011

Mechanics	2011-2021 Rating
Action Retrieval	7.49
Bingo	7.20
Force Commitment	7.71
Hot Potato	6.04
Legacy Game	8.19
Measurement Movement	7.30
Multiple Maps	7.62
Passed Action Token	8.80
Auction: Dutch Priority	7.08

4.3 Relationships between Success of Board Games and Their Attributes

The following heatmap illustrates rough relationships between pairs of different numeric attributes. Correlation values in heatmap conform to Pearson's coefficients:

$$\rho_{X,Y} = \frac{cov(X,Y)}{\sigma_X \sigma_Y} = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y}$$

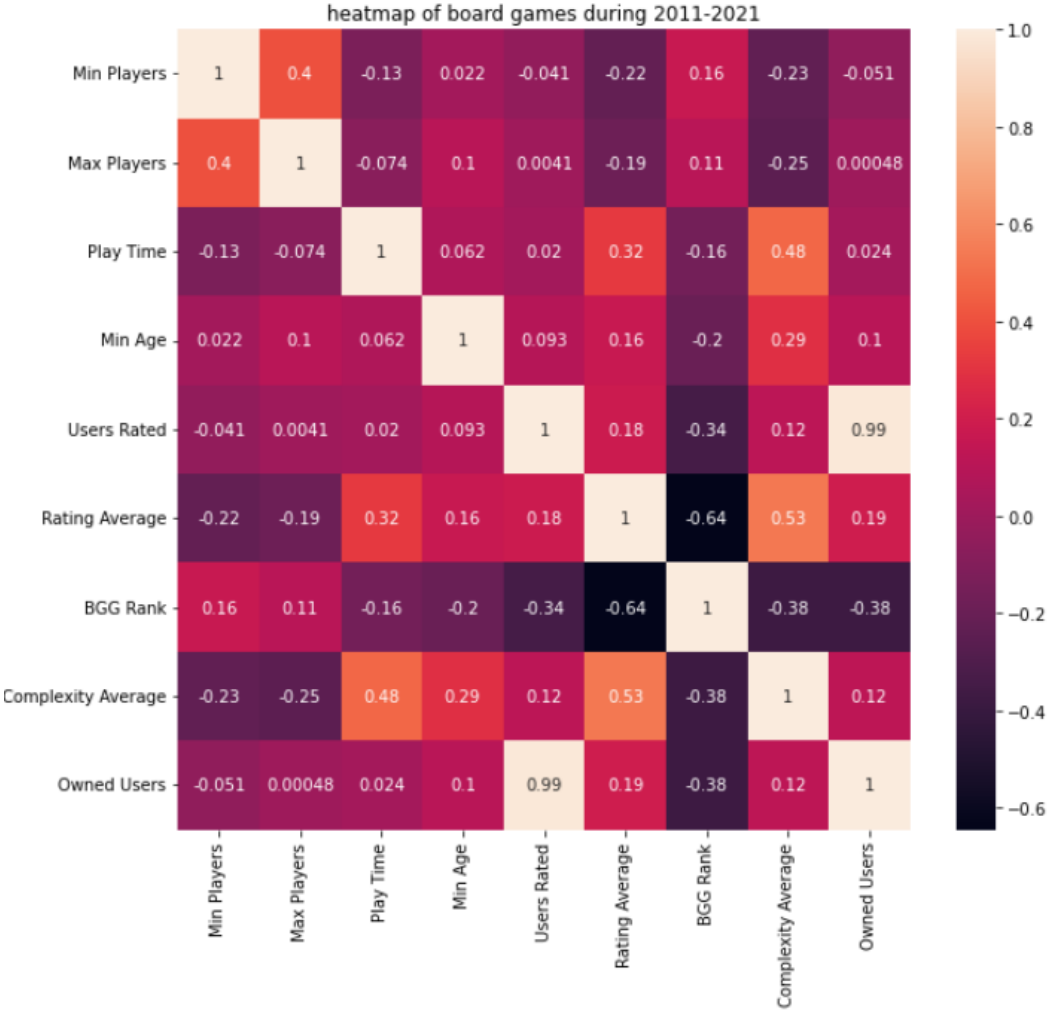


Figure 8. Heatmap of board games during 2011-2021

Some conclusions can be made according to Figure 8:

(1) *Users* Rated is completely relevant to *Owned Users*;

(2) *Min Age* is relevant to *Complexity Average*, which is rational as complex board games are commonly not oriented toward the low ages.

(3) Even with higher ratings, complex games don't have a significant effect on attracting board game buyers;

(4) Complexity of board games and Playtime have an apparent relationship with *Rating Average*, while these two factors also have close connections between each other.

As ratings of the board games are scored by the reviewer's posterior, both the rating and number of owned users will be considered factors that measure successes of a board game, the first one can be called the gameplay success, while the latter one is the commercial success. Conclusions (3) and (4) will be further discussed with the following detailed figures.

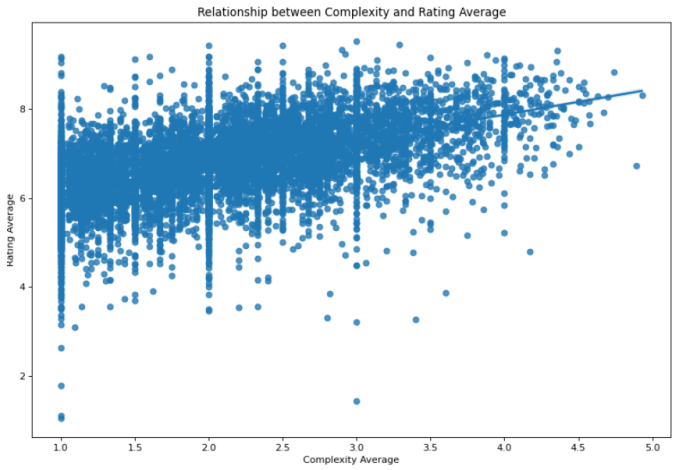


Figure 9. Relationship between Complexity and Rating Average

In the light of figure 9, complex board games do have higher ratings on average, but number of board games published decreases remarkably as the complexity level increases. Many high-rated board games also appeared with complexity averages of over 2.0 compared to difficult ones, and instead, it is the large quantity of low-rated games which lowers the *Rating Average*.

In conclusion, complex games surely have higher averages, but their high public praises tend to be more relevant to the 'by-product' high quality the game designers provided while advancing the difficulty.

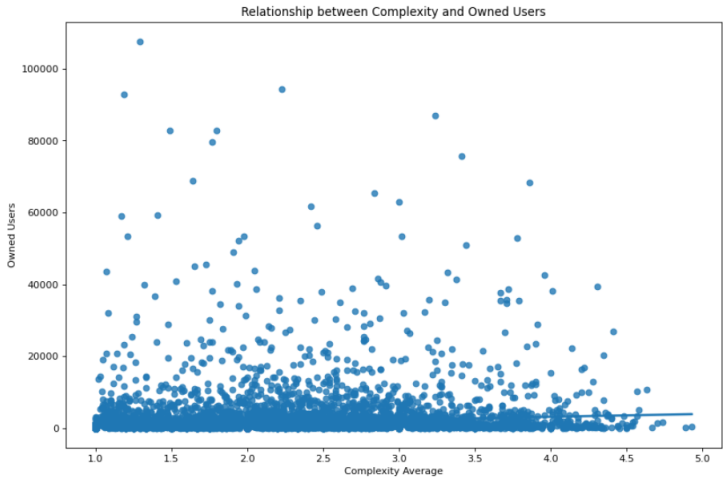


Figure 10. Relationship between Complexity and Owned Users

Figure 10 further proves that even leads to higher ratings, complexity may not become an attraction for users, as is seen that board games with complexities of 1 to 4.5 all have similar user structures.

Further, it can be inferred that for game publishers, developing lightweight board games might be more profitable and less time-consuming, thus becoming mainstream in the recent 10 years.

To see the complex influence of complexity and playtime which is presented in conclusion (4), a K-means++ Cluster Analysis methodology is used to classify board games with different attributes interval which were released between 2011 and 2021.

The purpose of the K-means++ algorithm is to select the centroid to minimize the inertia value within each cluster. The calculation method of inertial is as follows:

$$E = \sum_{i=1}^{k} \sum_{x \in C_i} ||x - \mu_i||_2^2$$

Steps of K-means++ methodology:

(1) Randomly select a sample from the dataset as the initial cluster center c_1;

(2) First, calculate the shortest distance $D(x)$ between each sample point and the current existing cluster center; Then calculate the probability $\frac{D(x)^2}{\sum_{x \in X} D(x)^2}$ that each sample is selected as the next cluster center. Finally, select the next cluster center according to the roulette method;

(3) Repeat step (2) until k cluster centers are selected;

(4) Calculate the Euclidean distance from other sample points to the center vector of each class, and divide it into the class with the closest distance;

$$\text{3D Euclidean distance: } d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}$$

(1) Update the center vector of each class;

(2) Determine whether the new class center vector has changed, if so, go to step (4). If the class center vector no longer changes, stop and output the clustering result.

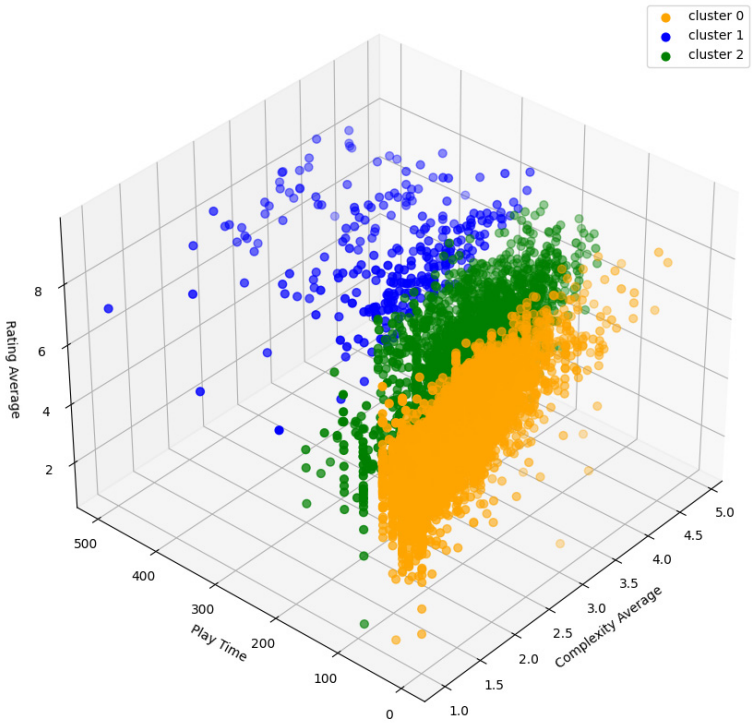


Figure 11. Complexity average

The 3 cluster' centers and counts are predicted as follows (x: *Play Time*, y: *Complexity Average*, z: *Rating* Average):

Table 6. Cluster Information

Cluster	Center	Count
Cluster 0	[34.26795693,1.73829767,6.54758212]	7337
Cluster 1	[113.65358284, 2.75953446,7.21018713]	2191
Cluster 2	[301.56630824,3.26154122 ,7.64189964]	279

In the light of figure 11, board games released from 2011 to 2021 are divided into 3 clusters, and the automatic clustering result and cluster centers in table 6 reveal that playtime plays a critical role in determining the classification and Rating of board games, with fast-playing games occupying 74.75% out of 9807 games and the most time-consuming ones taking 2.84%. While each of the three clusters renders a triangle on the y-z level, it also clearly shows that as the Complexity Average increasing, the quantity of board games drops extremely and at the same time, the average rating rises remarkably. Cluster centers in table 6 demonstrate the relationships between ratings, complexity, and playtime, with the apparent ascent on each attribute as the element quantity of each cluster decrease.

5. Conclusion

With a history of more than 5000 years, board games keep surviving and thriving throughout the year. From abstract games like chess and 'go' which were fashionable in ancient times to the 'GloomHaven' which obtains millions of fans around the world, the trend of board game popularity never stops to change.

By analyzing 20343 board games collected on the 'BoardGameGeek' website, the attributes and changes of the board games are visualized and summarized, in this analysis, historical changes in the board games are revealed. Showing that mechanics such as Dice Rolling, Hand Management and Set Collection and domains such as Wargame, and Strategy Game have been popular for decades, while in recent years, newcomers like Cooperative Games get a slice of the cake.

The highly difficult games successfully win praise from players, but the efforts their designers spent may not be worthy of the final profits. Instead, casual and delicate games might be more lucrative for publishers. Also, the result of the thousands of games released in the last ten years confirms the significant importance of the relationship between domains/mechanics and the games' successes, thus is worth further study of the upcoming trends of the board game market for game designers to succeed. With a view of the trend of board game published, it shows that the boosting of the board game market increases the number of failures of design. Publishers should stay more rigorous, to avoid losses in this pop market.

Further, the analysis may be access to help researchers make a better view of the evolution process of boardgames and dig into hidden relationships connecting with society, technology, or even international situation. The attributes of play time may relate to the duration of entertainment people enjoys, while the attributes of fashionable domains and mechanics reveal the ideological trends of board game players.

References

[1] Masukawa, Koichi. "The Origins of Board Games and Ancient Game Boards." Simulation and Gaming in the Network Society. Singapore: Springer Singapore, 2016. 3-11.

[2] Siegler R S, Ramani G B. Playing linear number board games—but not circular ones—improves low-income preschoolers' numerical understanding[J]. Journal of educational psychology, 2009, 101(3): 545.

[3] Paris T N S T, Yussof R L. Enhancing grammar using board game[J]. Procedia-Social and Behavioral Sciences, 2012, 68: 213-221.

[4] Bochennek K, Wittekindt B, Zimmermann S Y, et al. More than mere games: a review of card and board games for medical education[J]. Medical teacher, 2007, 29(9-10): 941-948.

[5] Samarasinghe D, Barlow M, Lakshika E, et al. A Data Driven Review of Board Game Design and Interactions of Their Mechanics[J]. IEEE Access, 2021, 9: 114051-114069.

[6] Dietrich T, Mulcahy R, Knox K. Gaming attribute preferences in social marketing programmes: Meaning matters more than rewards[J]. Journal of Social Marketing, 2018.