Analysis and Assessment of Changes in Game Production and Their Impact on Player Satisfaction

Abstract

This research focuses on the evolving trends in game production and their subsequent impact on player satisfaction by analyzing a comprehensive dataset from the Steam Store. The study examines critical metrics such as game ratings, pricing patterns, platform support, and review behaviors over time. Results indicate that while positive ratings are becoming more prevalent, overall game ratings have declined, possibly due to increasing player expectations. The analysis also highlights significant trends, including the rise of free-to-play games, the lack of correlation between reviews and satisfaction, and platform-independent ratings. These insights emphasize the need for innovation and player-focused development in the dynamic gaming industry.

Introduction

The gaming industry is an ever-evolving domain, shaped by technological advancements, shifting player expectations, and diverse production strategies. Digital distribution platforms such as Steam have revolutionized how games are developed, marketed, and consumed. Launched by Valve Corporation in 2003, Steam has grown to become a global hub for video games, hosting over 400 million titles by 2023. Its community-driven approach allows players to provide feedback, influencing game ratings and driving industry trends.

This research investigates how these shifts in game production affect player satisfaction, with a specific focus on metrics such as game ratings, pricing, platform accessibility, and user reviews. Using the "Game Recommendations on Steam" dataset, the study seeks to uncover patterns and correlations in player feedback and industry practices over time.

Through data cleaning, visualization, and statistical analysis, this paper provides insights into the industry's dynamic nature, highlighting challenges and opportunities for developers. By examining temporal trends and emerging patterns, this research contributes to understanding how production decisions align with player expectations and satisfaction.

Literature Review

Valve Corporation developed the digital distribution platform and store known as Steam. It began as a software client in September 2003, providing automated game updates for Valve's games, and expanded to distribute third-party titles in late 2005. [2] Steam is widely regarded as one of the world's most popular online gaming platforms and distributors. In 2023, Steam was estimated to sell over 400 million games overall. [3]

"During game development, a common uncertainty revolves around the emotional reactions players might exhibit when experiencing newly released games. Understanding emotions and experiences of video game players is complex. As a first step, players can share their experiences of video games by writing reviews. By investigating these reviews, the emotions, experiences, concerns, and opinions of players can be understood." [4]

"An intriguing aspect of Steam community is its evolution over time." [5]

Materials and Methodology

"Game Recommendations on Steam" dataset [1] has been analysed through different approaches. The dataset contains over 41 million reviews from the Steam Store.
A common way for users to express their satisfaction with a game is by leaving a review. On Steam, after purchasing and playing a game, a user can write a review to describe what they liked and disliked about the game, and select on whether they would recommend the game or not. That way, all reviews can be either positive or negative. Steam will then take the ratio of positive and negative reviews given to a game, to determine the game's overall rating, which can be overwhelmingly positive, mostly positive, very positive, positive, mixed, negative, very negative, mostly negative and overwhelmingly negative.
The dataset used for this research contains the following columns: 'app_id', 'title', 'date_release', 'win', 'mac', 'linux', 'rating', 'positive_ratio', 'user_reviews', 'price_final', 'price_original', 'discount', and 'steam_deck'.
The 'app_id' column represents the native product ID on Steam, 'title' column contains the title of the game, and 'date_release' is the product's release date. Columns 'win', 'mac', 'linux' and 'steam_deck' contain boolean values which represent whether the game is supported on each operating system/platform, where True means the game is supported and False means that it is not. The column 'rating' represents the product's rating category (overwhelmingly positive, mostly positive, etc.), the 'positive_ratio' contains the ratio of positive feedback, and 'user_reviews' is the amount of user reviews available on the Steam page of the game. [1]

Data was analysed using the Python programming language. After importing all the necessary libraries and loading the dataset, the research started with data cleaning and preprocessing, through which it was discovered that the dataset has 50872 games, with 13 columns, and no null values. Out of the 13 columns, 4 of them contain boolean data, 3 contain floats, 3 integers and 3 object data. We have also discovered that 'positive_ratio' had the highest rated game with 100% positive feedback and the lowest rated game which had 0% positive feedback. On average games have 77% positive feedback. 'price_final' shows us that the most expensive game reached nearly 300 USD, whereas the cheapest game was free. 25% of all games cost under 1 USD, and 50% of games cost under 5 USD, but the average game price was 8.6 USD. The column 'discount' shows us that at least 75% of games had no discount applied, but the game with the highest discount was lowered by 90%. That said, the average discount was 5.6% The dataset did not contain any duplicate values either.

In order to easily represent the 'rating' values on a graph, the ratings have been converted from categorical to numerical data, as can be seen in the figure below.

```python
# Defining the mapping dictionary
rating_mapping = {
    'Overwhelmingly Negative': 0,
    'Mostly Negative': 1,
    'Very Negative': 2,
    'Negative': 3,
    'Mixed': 4,
    'Positive': 5,
    'Very Positive': 6,
    'Mostly Positive': 7,
    'Overwhelmingly Positive': 8
}

# Applying the mapping to the 'rating' column
games['rating_numeric'] = games['rating'].map(rating_mapping)
```

Figure 1. Converting Ratings From Categorical to Numerical

There was no longer any need for the 'rating' column, so it got removed from the dataset.
Once again, for easier data visualization and analysis, the 'date_release' column has been converted from object to datetime, as shown in the figure below.

```python
#Converting date_release from object type to datetime

games['date_release'] = pd.to_datetime(games['date_release'])
```

Figure 2. Converting Release Date From Object to Datetime

To get a better understanding of the data, we performed data visualization.
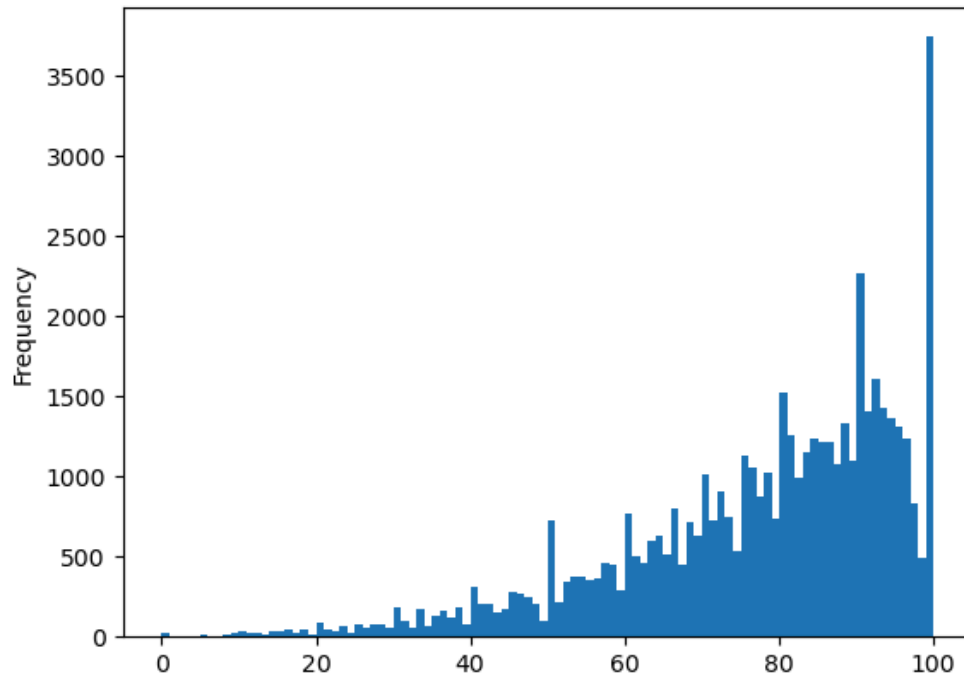Figure 3 shows a graph representing the positive ratio of game reviews.

Figure 3. Positive Ratio Graph

On the other hand, in Figure 4, which represents the graph for game ratings, it can be seen that most games belong in the 'Positive' category, forming a sort of a bell-curve. That could be explained by Steam's categorization guides, according to which a game can be categorized as 'Overwhelmingly Positive' not only if it has a high positive ratio, but also if it exceeds a threshold for the minimum number of reviews. Same applies for each category.
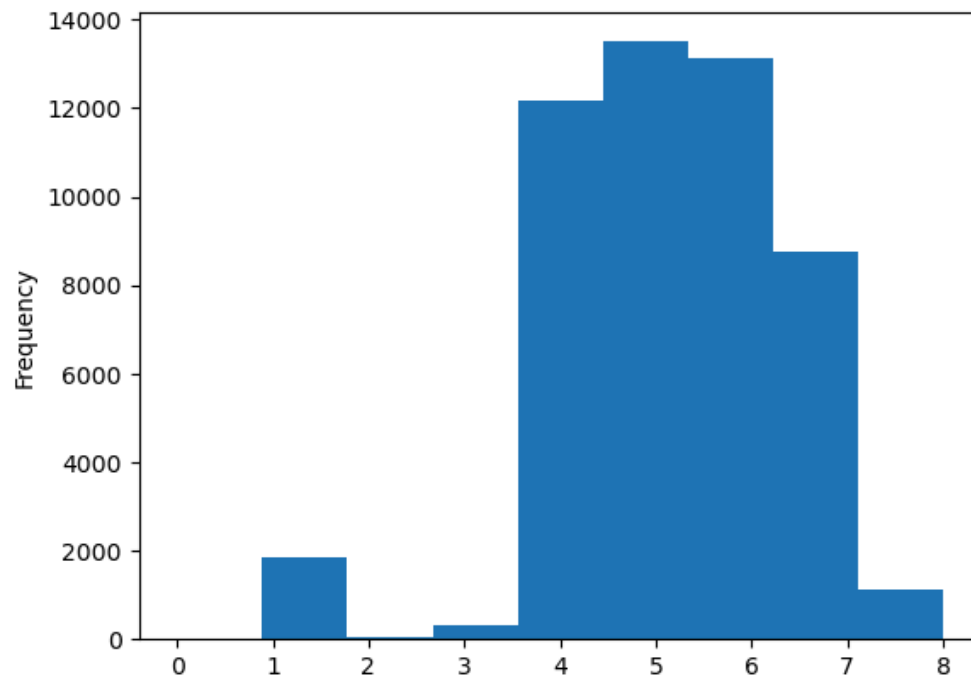

Figure 4. Game Ratings Graph

After seeing that most games have Windows and Steam Deck support, but only a fraction of them had Mac and Linux support (Figure 5), the next question was does that affect the average game rating of the game, and as can be seen in Figure 6, it does not. With that we can conclude that no platform lacks or excels in its respective game development.
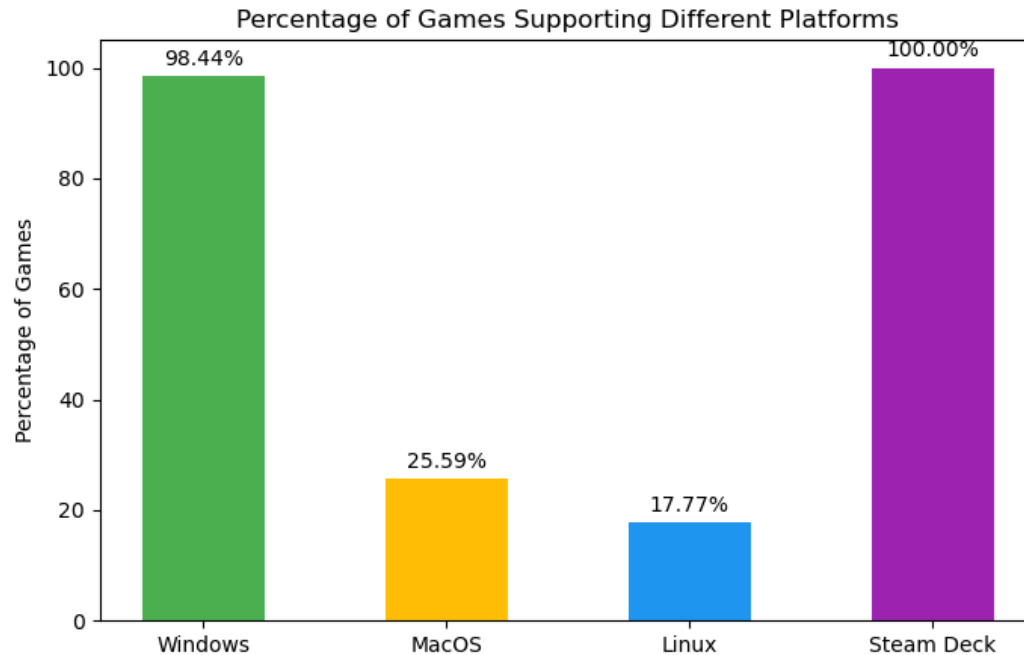

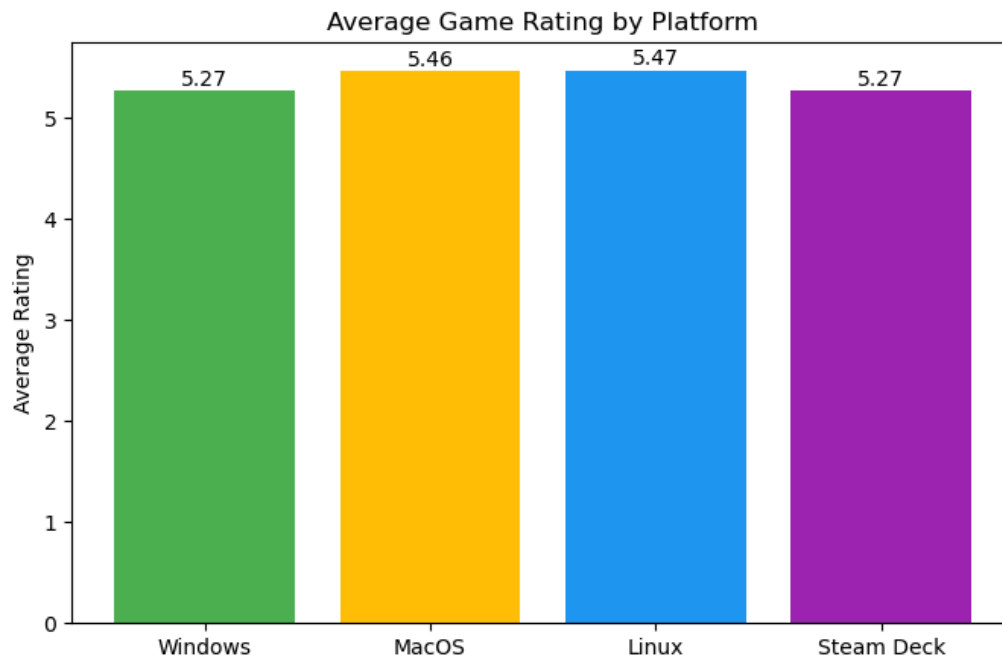Figure 5. Percentage of Games Supporting Different Platforms

Figure 6. Average Game Rating by Platform

With the plot represented in FIgure 7 we can see how the game ratings have become lower over time, with an unusual drop in 2003. This could either mean that the game development got worse, or, the more likely option, that the players became stricter and more rigorous with their reviews. With the increase in game players there is an increase in people who review the game (Figure 8), which may be positive or negative. To test the theory the plot shown in Figure 9 was created, representing both the game ratings and the number of ratings over time. Although there is no precise overlap between the two, we can see that as the number of ratings was increasing, the average game rating was decreasing. We can also assume that, with more games being produced over time, people refined their likings of specific games and have more narrow preferences than before, but there is no way to actually prove this theory. However, to further test whether the number of reviews has affected the ratio of positive reviews, the graph has been shown in Figure 10, and in Figure 11 we can notice the Pearson correlation, according to which there is no significant correlation between the number of reviews and the ratio of positive reviews.
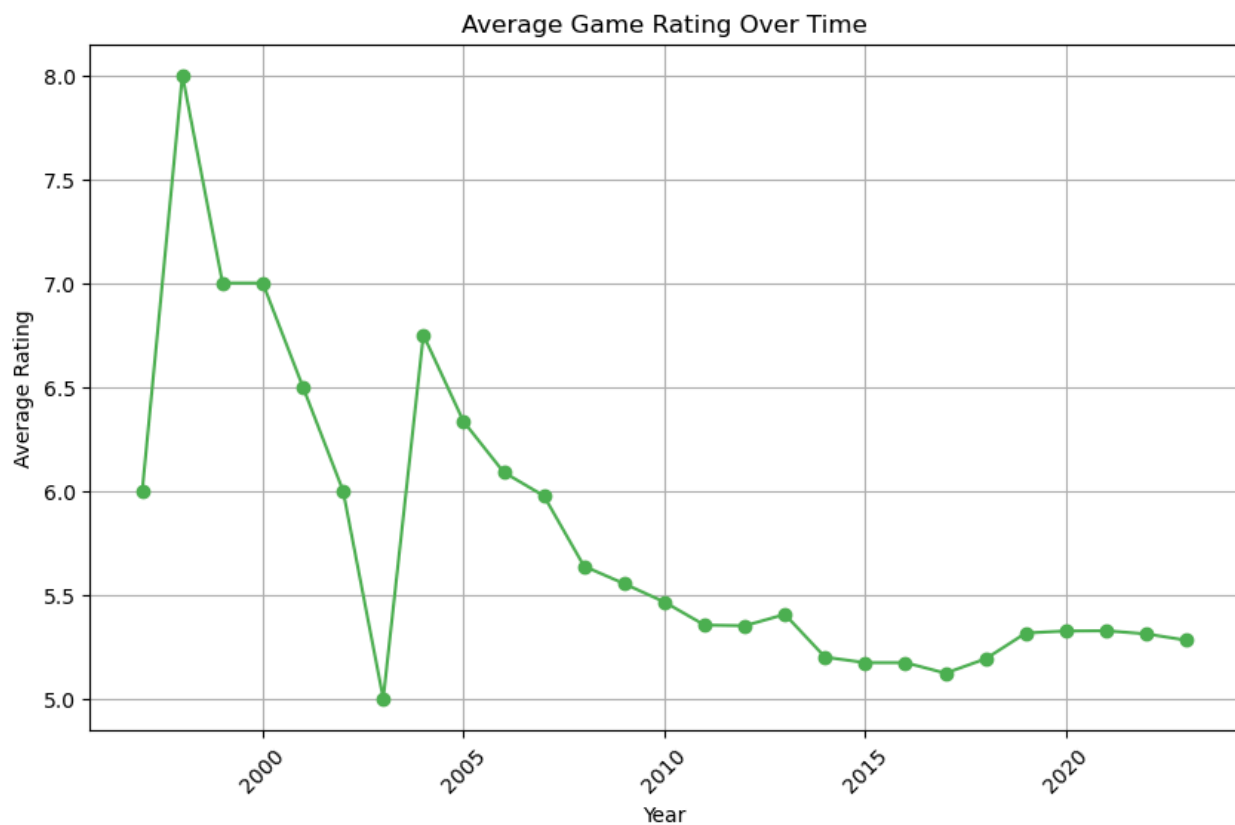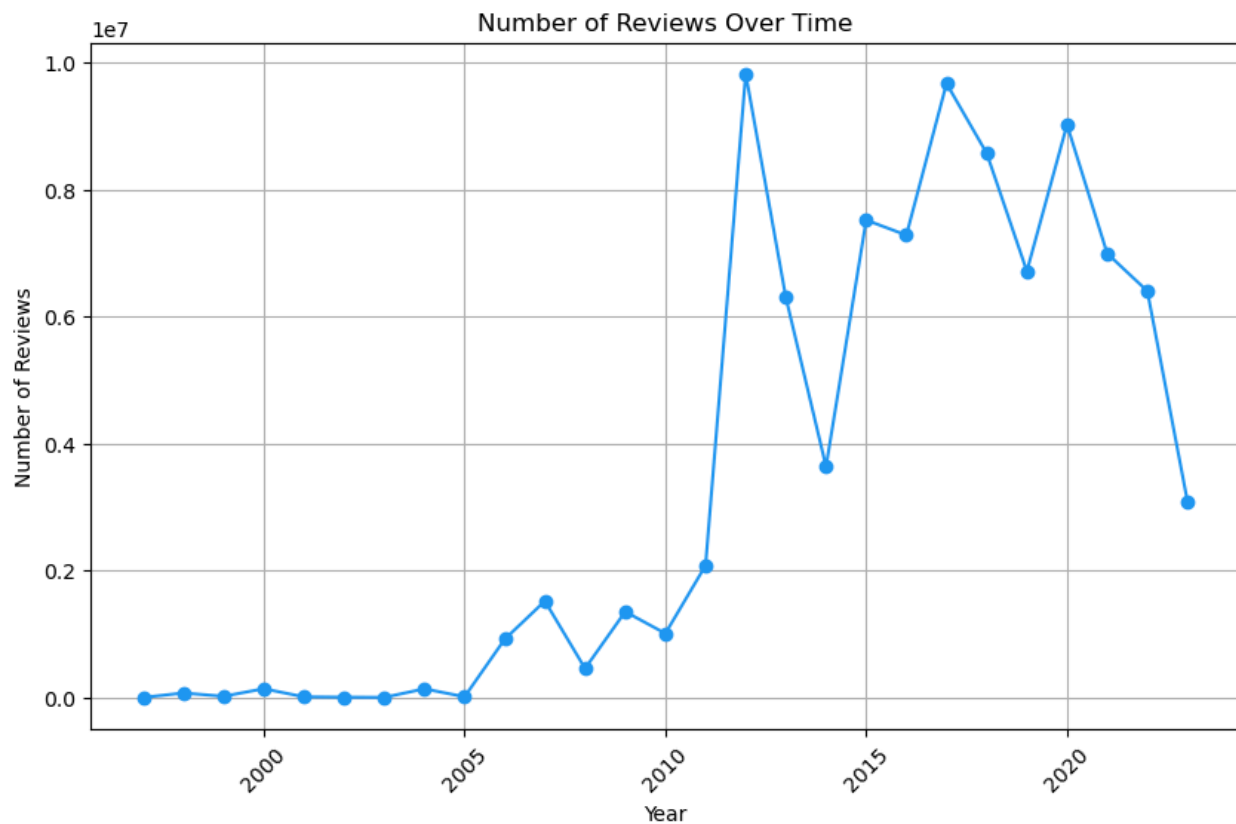

Figure 7. Average Game Rating Over Time
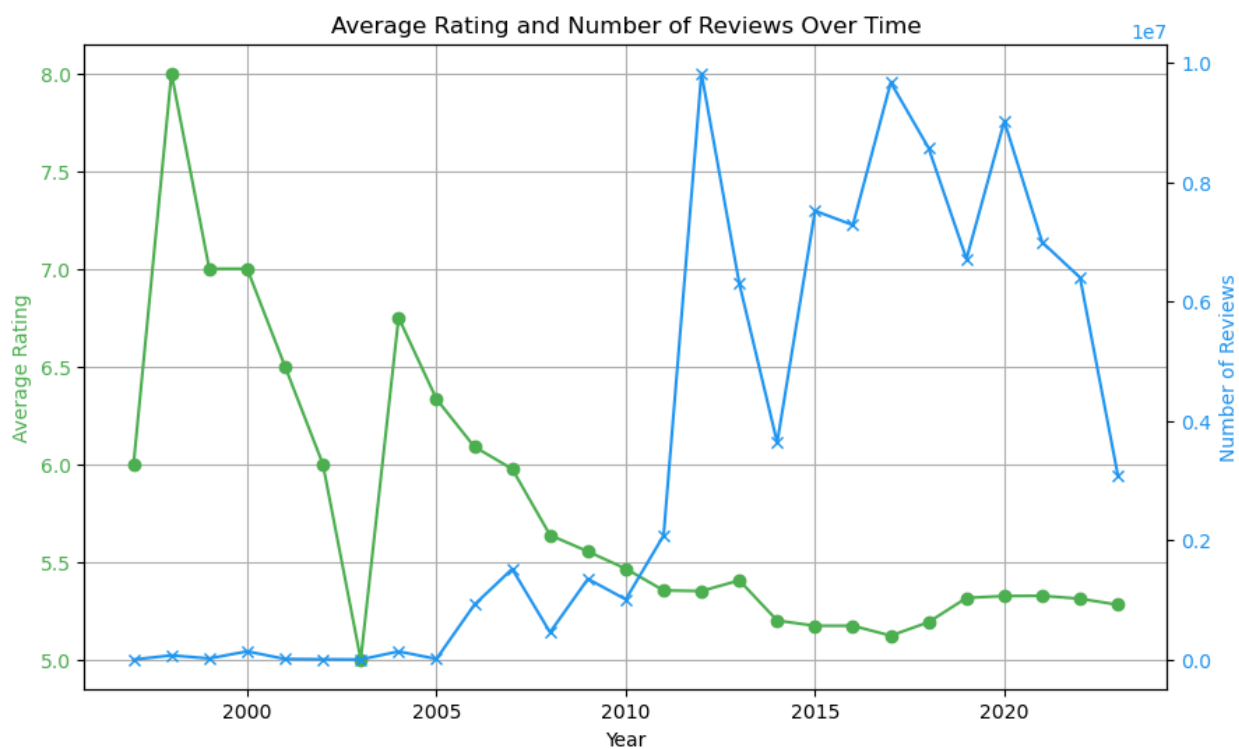
Figure 8. Number of Reviews Over Time

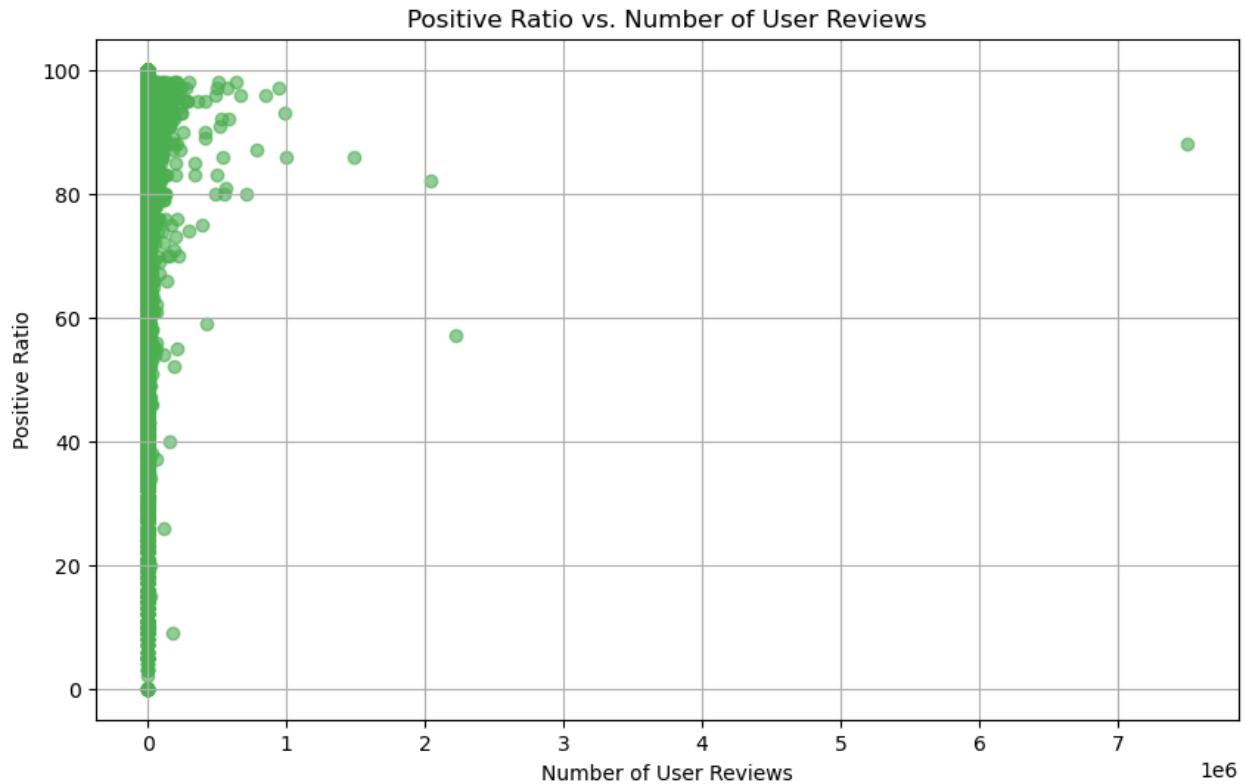Figure 9. Average Rating and Number of Reviews Over Time

Figure 10. Positive Ratio vs. Number of User Reviews

```
# Calculate Pearson correlation between user_reviews and positive_ratio
correlation = games[['user_reviews', 'positive_ratio']].corr().iloc[0, 1]

print(f"Pearson Correlation Coefficient: {correlation}")

Pearson Correlation Coefficient: 0.021067961210372608
```

Figure 11. Pearson Correlation

Figure 12 shows us that, on average, the games were cheapest in 2005, and are currently the most expensive they ever were. However, we can also see in Figure 13 that there has been an increase in the number of free games as well, reaching its high in 2023 and decreasing since.
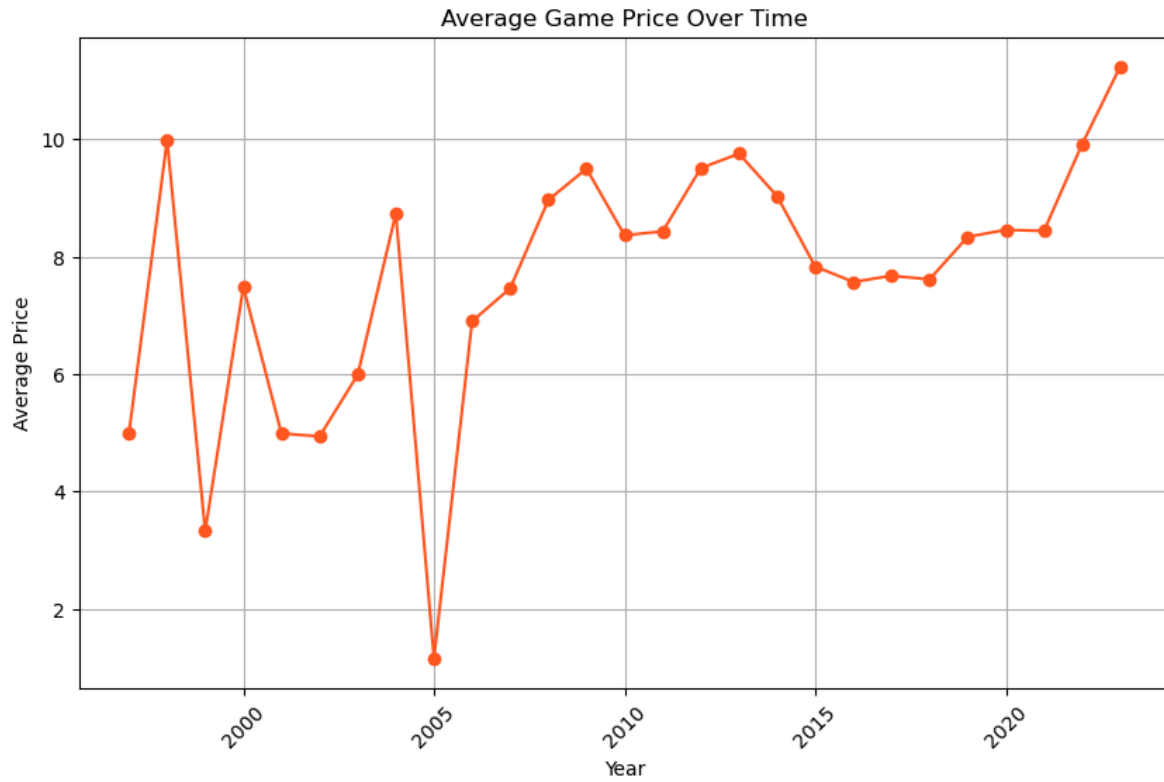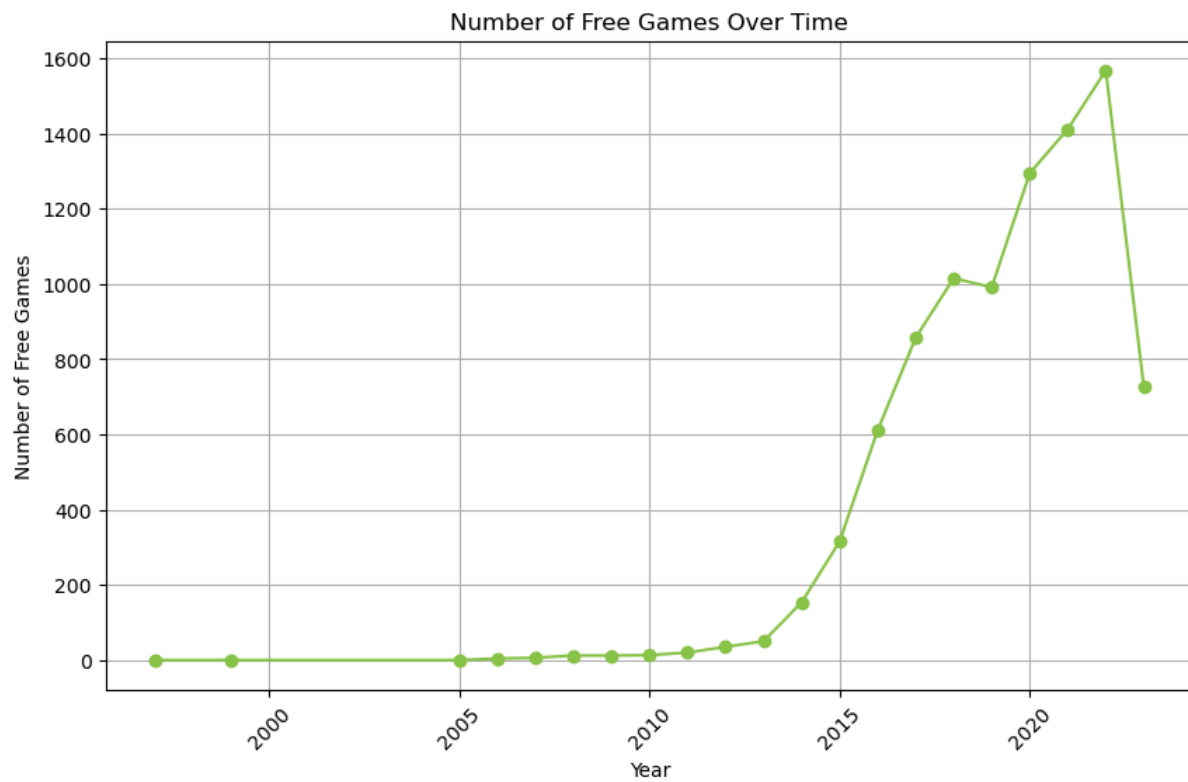
Figure 12 Average Game Price Over Time

Figure 13 Number of Free Games Over Time

Results and Discussion

Findings:
- The ratio of positively rated games increases (there are more games with high positive ratio than games with low positive ratio)
- Majority of games are low priced (most of them are free)
- Game rating is distributed almost as a bell curve, where most games have a rating of 5, then 6, then 4, then 7, then 1, 8, 3, 2, and 0
- All games are supported on Steam Deck, most on Windows, some on Mac and even less on Linux
- Game platform does not affect the game rating
- Game rating is decreasing over time, with newly released games having lower ratings than the older games
- The number of reviews increases over time, with newer games having more reviews than the older games
- Number of reviews and the rating do not correlate
- Prices of games are increasing over time
- The number of free games also increases over time
- There is no correlation between the number of reviews and ratio of positive reviews
- The 5 highest positive ratio games with the most reviews are: Aokana - Four Rhythms Across the Blue - EXTRA2, Endless Monday: Dreams and Deadlines, 祈風 Inorikaze, South Scrimshaw Part One, UOS Prototype
- The 5 lowest positive ratio games with the most reviews are: Fray: Reloaded Edition, Heroes in the Sky-Origin, Lucky Night: Poker Games, Arachnophobia, Christmas Carol
- The 5 highest rated games with most reviews are: Counter-Strike: Global Offensive, Dota 2, Grand Theft Auto V, Tom Clancy's Rainbow Six® Siege, Team Fortress 2
- The 5 lowest rated games with the most reviews are: Overwatch® 2, War of the Three Kingdoms, Expansion - Europa Universalis IV: Leviathan, Spacebase DF-9, Flatout 3: Chaos & Destruction
- Top 10 most reviewed games are: Counter-Strike: Global Offensive, PUBG: BATTLEGROUNDS, Dota 2, Grand Theft Auto V, Tom Clancy's Rainbow Six® Siege, Team Fortress 2, Terraria, Garry's Mod, Rust, Apex Legends™

Positive Trends in Ratings: The distribution of positive ratings shows an increase in the ratio of highly rated games, indicating an overall trend toward better-received games. However, the bell curve distribution of ratings suggests that most games fall within an average rating range, with fewer games achieving either very high or very low ratings.

Pricing Patterns: The majority of games are either low-priced or free, making gaming increasingly accessible. Despite this, the prices of paid games are rising over time, reflecting shifts in production costs, market demand, or monetization strategies.

Increase in Free Games: The number of free-to-play games is increasing over time, aligning with industry trends toward alternative revenue models like microtransactions or subscriptions.

Platform Accessibility: Steam Deck supports all games analyzed, while Windows remains the dominant platform. MacOS and Linux support are significantly lower, reflecting a continued preference among developers for platforms with larger user bases. Platform choice does not appear to influence game ratings significantly.

Temporal Trends in Ratings and Reviews: Game ratings are declining over time, with newer games generally receiving lower ratings than older titles. This decline could be attributed to increased player expectations or oversaturation of similar game types in the market. At the same time, the number of reviews per game is increasing, reflecting a growing player base or enhanced community engagement. However, no significant correlation was noticed between the number of reviews and positive review ratios, nor between the number of reviews and overall ratings. This suggests that high visibility or large player bases do not necessarily lead to better or worse game satisfaction.

Game Popularity Highlights: Analysis of the most-reviewed games reveals that enduringly popular titles such as "Counter-Strike: Global Offensive" and "Dota 2" remain dominant. Similarly, lists of the highest and lowest-rated games highlight the diversity in quality and reception, emphasizing the critical role of production values, gameplay innovation, and community feedback.

Conclusion

This research analyzed changes in game production and their impact on player satisfaction by examining several game characteristics such as pricing, platform support, rating, and user reviews. Overall, the data show a dynamic industry with constantly changing trends in production, pricing, and player participation. While accessibility and community interaction are growing, maintaining high quality and meeting the expectations of an increasingly demanding player base remains a challenge. These findings show the importance of continuous innovation and focus on players for long-term success in the gaming industry.

References

1. Kozyriev, A. (2024). Game Recommendations on Steam. Kaggle.com. https://www.kaggle.com/datasets/antonkozyriev/game-recommendations-on-steam/data
2. Wikipedia Contributors. (2019, October 21). Steam (service). Wikipedia; Wikimedia Foundation. https://en.wikipedia.org/wiki/Steam_(service)
3. Statista (2024, January 10). Topic: Steam. Statista. https://www.statista.com/topics/4282/steam/#topicOverview
4. Guzsvinecz, T., & Szűcs, J. (2023). Length and sentiment analysis of reviews about top-level video game genres on the steam platform. Computers in Human Behavior, 149, 107955.
5. Becker, R., Chernihov, Y., Shavitt, Y., & Zilberman, N. (2012, November). An analysis of the steam community network evolution. In 2012 IEEE 27th Convention of Electrical and Electronics Engineers in Israel (pp. 1-5). IEEE.