PROJECT REPORT



Analyzing Genre Trends in Video Games

**Master of Science in Business Analytics**

Course: MSBX 5420 Unstructured & Distributed Data

Prof. Zhiyi Wang

**Group 13 Team Members**

Akshaj Khirwadkar – akkh8908@colorado.edu

Anubhav Dubey – andu7792@colorado.edu

Jake Watsky – jawa0071@colorado.edu

Raghavendra Pariti - vepa3976@colorado.edu

Table of Contents

Abstract.........................................................................................................................................3

Background & Motivation............................................................................................................4

Background.........................................................................................................................4

Motivation...........................................................................................................................4

Dataset & Analysis Methods.........................................................................................................5

Dataset Description.............................................................................................................5

Source.................................................................................................................................5

Data Preparation and Analysis Methodology.....................................................................6

Game Genre Overview.......................................................................................................6

Results & Insights.........................................................................................................................8

Top Games Analyses..........................................................................................................8

Review Score vs. Genre....................................................................................................10

Revenue vs. Genre............................................................................................................12

Ownership vs. Genre.........................................................................................................14

Exploring Game Data With TF-IDF and Cosine Similarity.......................................................15

Functional Blocks.............................................................................................................15

Implications & Strategic Recommendations...............................................................................17

Future Project Enhancements......................................................................................................20

Conclusion...................................................................................................................................23

References...................................................................................................................................24

Appendix.....................................................................................................................................25

Abstract

The project explores various aspects of the history of video game genres through an analysis of an extensive dataset that includes around 30,000 titles produced between 2016 and 2024. The main sources of information include sales data and game metadata, as well as aggregated user review scores from the most significant gaming platforms. The approach is based on advanced text analytics which utilizes cosine similarity and TF-IDF vectorization to categorize games into genres and subgenres, allowing for a more in-depth investigation of consumer preferences and market trends.

The analytical framework applied Python and associated libraries for robust data manipulation, enabling the investigation of genre popularity, sales dynamics, and critical reception. Our key analytical efforts include:

* Quantifying shifts in genre prevalence and introducing new hybrid genres.
* Correlating game sales with review scores to assess the impact of game quality on commercial success.
* Creating visualizations to compare the rise and fall of genre popularity and market performance across different periods.

Findings highlight a marked shift towards hybrid genres and a nuanced understanding of the relationship between critical acclaim and financial success, especially in genres like action and adventure. The study also surfaces actionable insights for game developers and marketers, emphasizing the strategic importance of quality and innovation in capturing consumer interest in a competitive market. This report not only enriches academic discourse but also offers practical implications for industry stakeholders, aiming to enhance strategic planning and development in the video game industry.

**Background & Motivation**

Background

Over the past few decades, the video game industry has changed dramatically, emerging as one of the most enticing and dynamic segments of the entertainment industry. The industry evolved from initially prioritizing basic gaming mechanics to developing detailed, eye-catching experiences that appeal to a wide range of audiences worldwide. The expansion of gaming platforms and digital game distribution have made games more accessible, which has led to an increase in the amount and variety of data about player involvement, game genres, and market trends. As a result of this development, video games are now studied both academically and commercially, with everyone trying to comprehend the dynamics of the industry and how consumers behave.

### Motivation

The motivation for our project was to comprehend the vast amount of data produced by the video game industry. Understanding these changes can be crucial for developers, marketers, and strategists working in the gaming industry as new genres appear and outdated one's change. This research is especially relevant considering how quickly game design and customer preferences are evolving due to changing cultural trends and technology breakthroughs. Additionally, there was a major desire for us to research into the relationship between user reviews and ratings and critical and commercial success, given the growing significance of these aspects in purchase decisions. By examining these factors, this research aims to uncover patterns and trends that can inform future game development and marketing strategies, ultimately contributing to sustained growth and innovation within the industry.

**The Dataset & Analysis Methods**

Dataset Description

For this research, we utilized a detailed video game dataset derived from the sales, ratings, and user reviews of around 30,000 video games across various platforms, spanning a period from 2016 to 2023. This comprehensive dataset is curated to include:

* Sales data across multiple regions, allowing for a global analysis of market trends.
* User ratings and reviews from prominent gaming websites, providing insights into player satisfaction and game reception.
* Tags and Keywords: Genre classifications and platform details for each game, which help in understanding the popularity and market dynamics of various game types.
* Release dates and publisher information, facilitating a temporal analysis of market shifts and the impact of different publishers in the industry.

This dataset enables a multi-dimensional analysis of the video game industry, focusing on trends in game sales, the relationship between user ratings and commercial success, and the evolution of game genres over the past two decades.

Source

* Steam API
* Kaggle

Data Preparation and Analysis Methodology

* **Dataset Segmentation by Genre**

Segmented the dataset to classify games into distinct genre categories for focused analysis.

* **Title Data Standardization**

Cleansed the 'Titles' column by removing all special characters to ensure data consistency across the dataset.

* **Data Type Conversion**

Transformed 'Review Scores' and 'Revenue Estimated' columns to numerical formats to enable quantitative analysis.

* **Data Cleansing**

Eliminated rows containing NaN values and duplicates to maintain data integrity and accuracy in the dataset.

* **Ownership Estimation through Review-Based Multiplier**

Applied the Boxleiter method leveraging the Review to Sale Multiplier along with the annual total number of reviews to estimate the ownership numbers for games categorized by genre.

Game Genre Overview

* **Battle Royale**

Competitive survival games where players are dropped into a shrinking play area and must be the last one standing.

* **Multiplayer**

Games designed for or that support multiple players to play simultaneously, either cooperatively or competitively.

* **Role Playing Games (RPG)**

Games where players assume the roles of characters in a fictional setting and take control over many of their actions.

* **Racing**

Games that simulate racing competitions, whether with cars, motorcycles, or other vehicles, often emphasizing speed and vehicle control.

* **Strategy**

Games that emphasize strategic, tactical, and sometimes logistical challenges, often with resource management and long-term planning.

* **Sports**

Games that simulate the playing of traditional physical sports, such as soccer, basketball, or golf.

* **Free to Play**

Games that are free to download and play, often monetized through in-game purchases or ads.

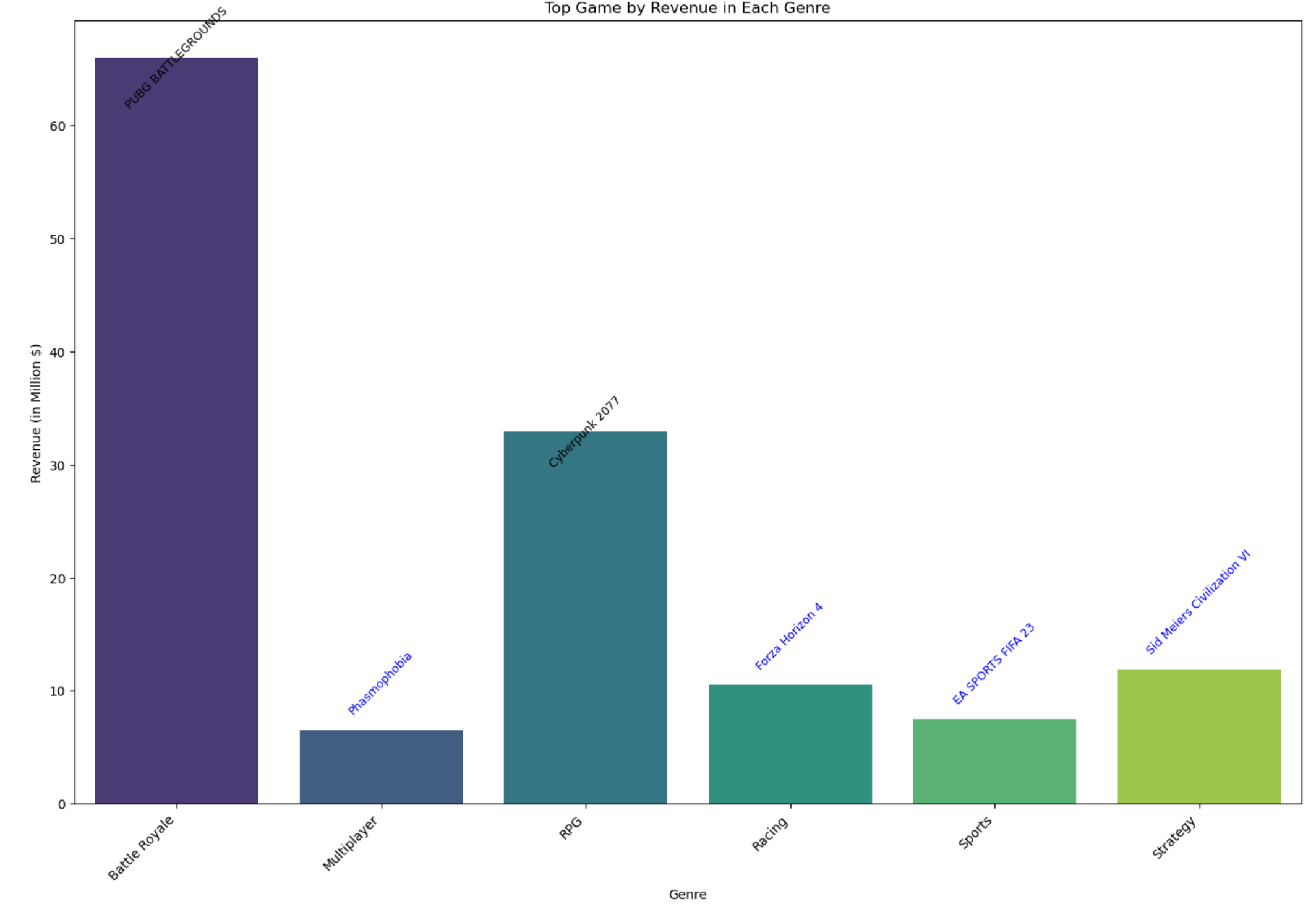
* **Paid**

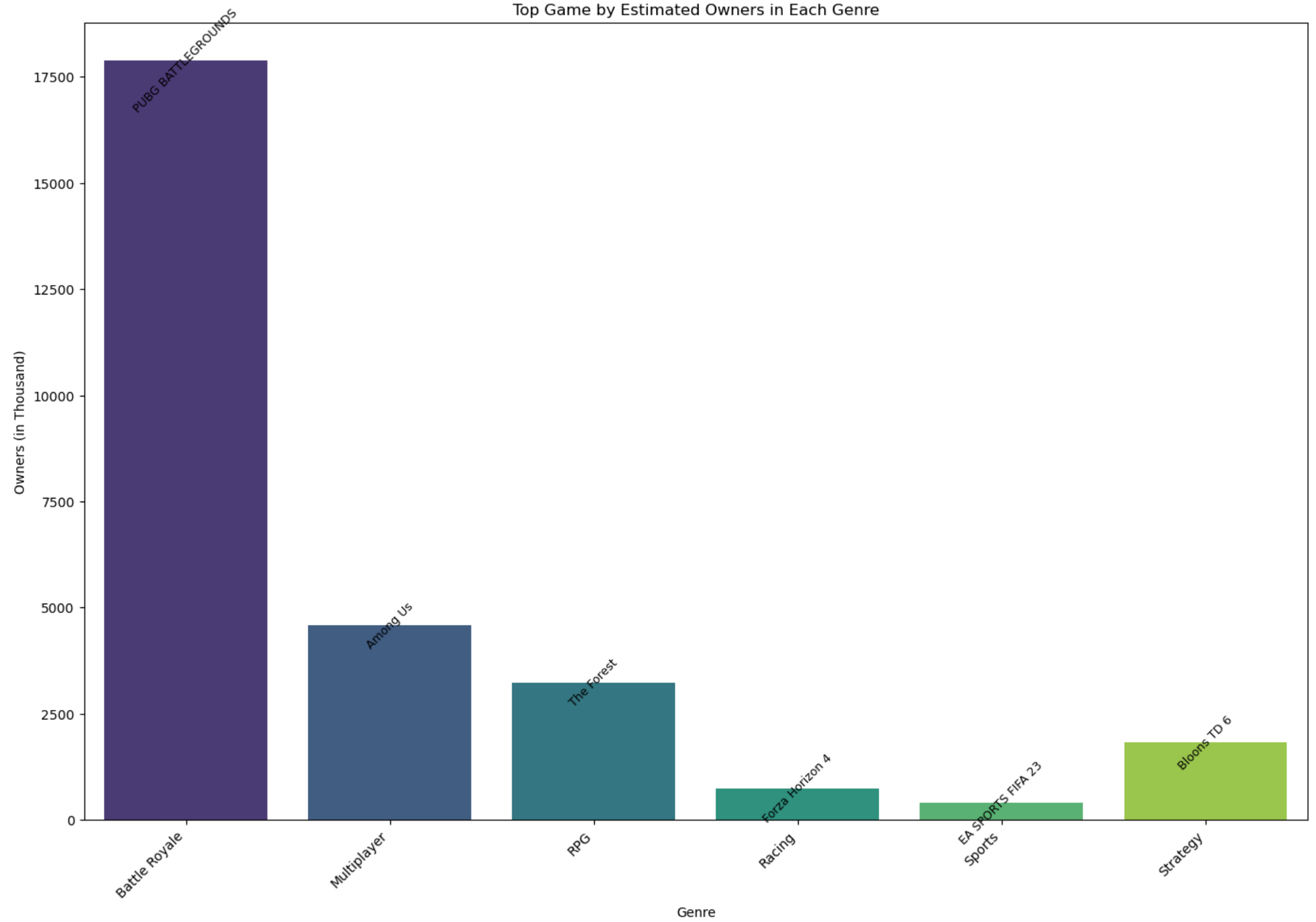
Games that require a one-time purchase to download and play, with no initial cost barrier.



**Results & Insights**

Top Games Analyses



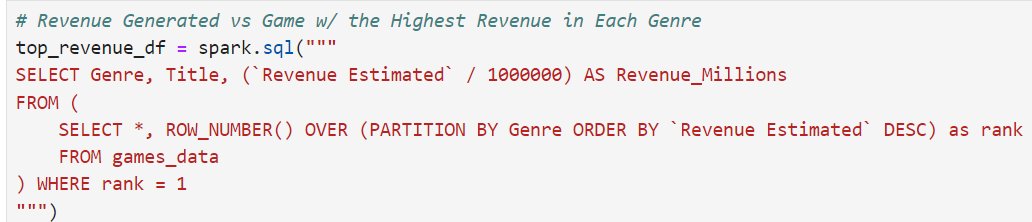


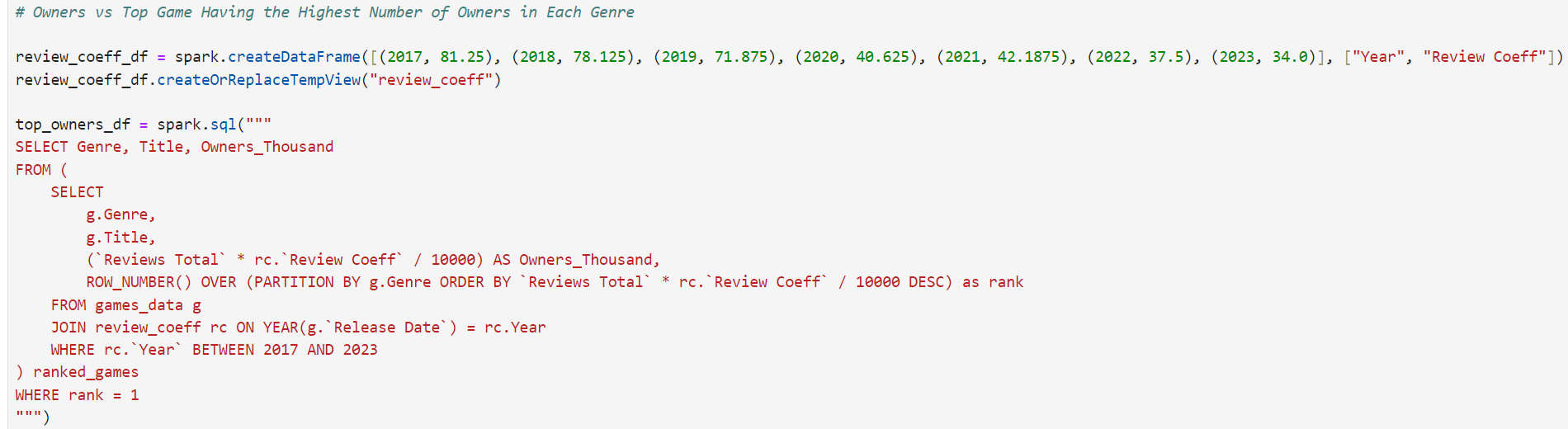
1. **PUBG: Dominance in the Gaming Market**

Recent analyses reveal that 'PUBG: Battle Royale' leads the gaming industry with the highest revenue among all genres, significantly outpacing its competitors. The game boasts an impressive ownership total of approximately 17.5 million. This strong performance underscores its market dominance and widespread appeal.

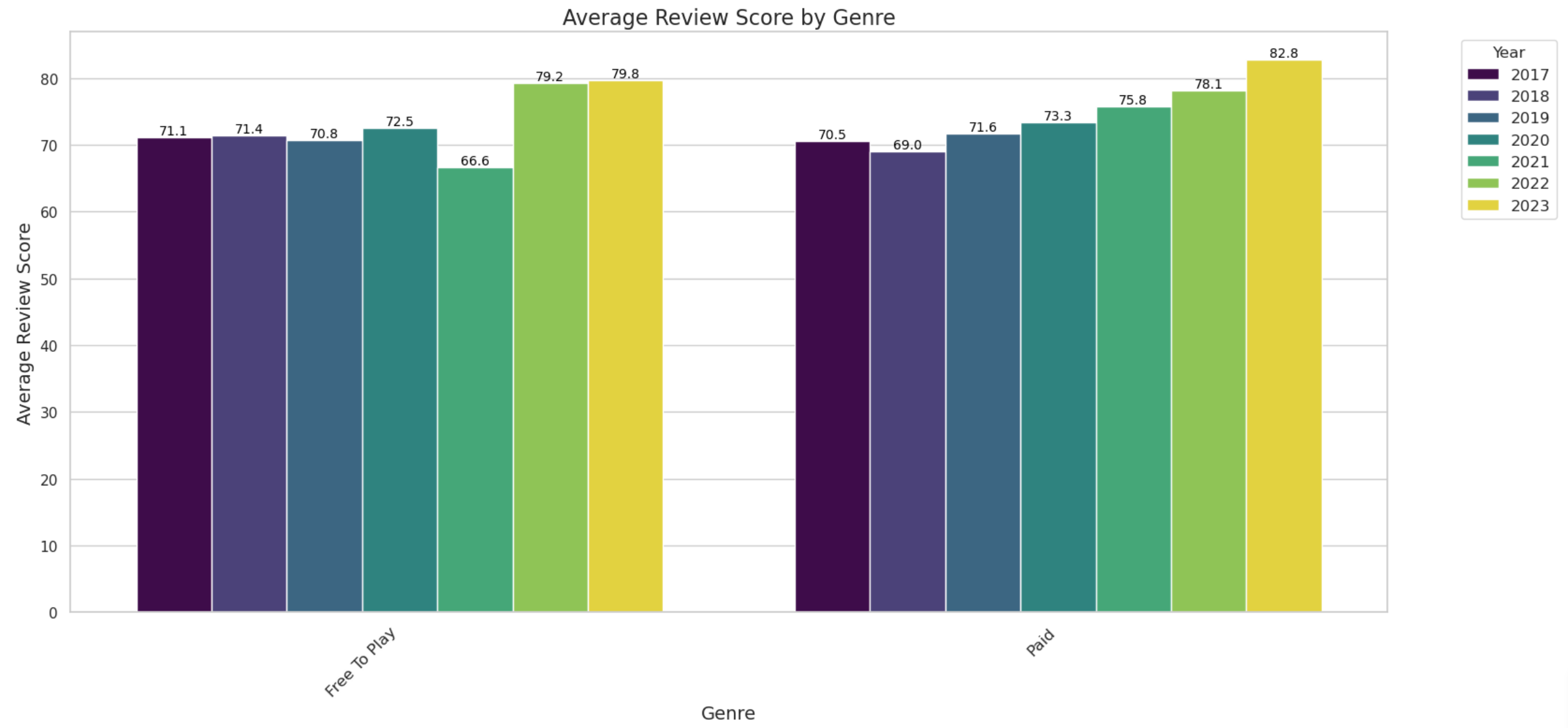
1. **Cyberpunk: A Leading RPG**

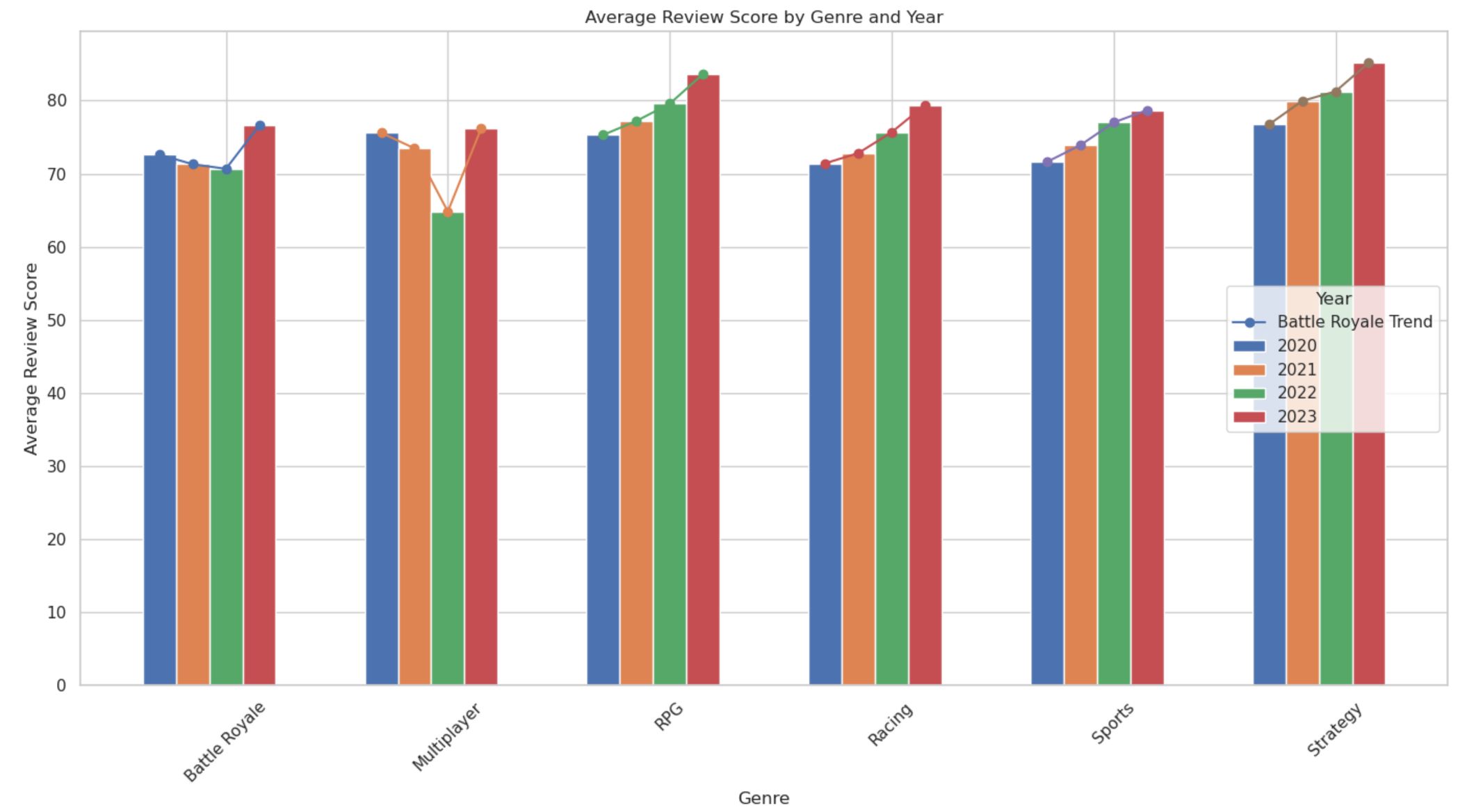
In the role-playing game (RPG) category, 'Cyberpunk' emerges as the top-selling title. Despite its relatively modest ownership figures, it secured the second position in revenue generation. This suggests that RPGs hold considerable potential for revenue, indicating a promising future for this genre in the gaming industry.





Review Score vs. Genre Trends





1. **Overall Increase in Scores Over Time**:

There appears to be a general trend of increasing review scores over time across all genres. This could suggest either an improvement in game quality, a shift in reviewing patterns, or changes in consumer expectations and preferences.

1. **Paid vs. Free to Play**:

The Paid genre consistently outperforms Free to Play in terms of average review scores. This might indicate that consumers perceive a higher value in paid games, which could be due to better production quality, content, or absence of in-game purchases which are often found in free-to-play models.

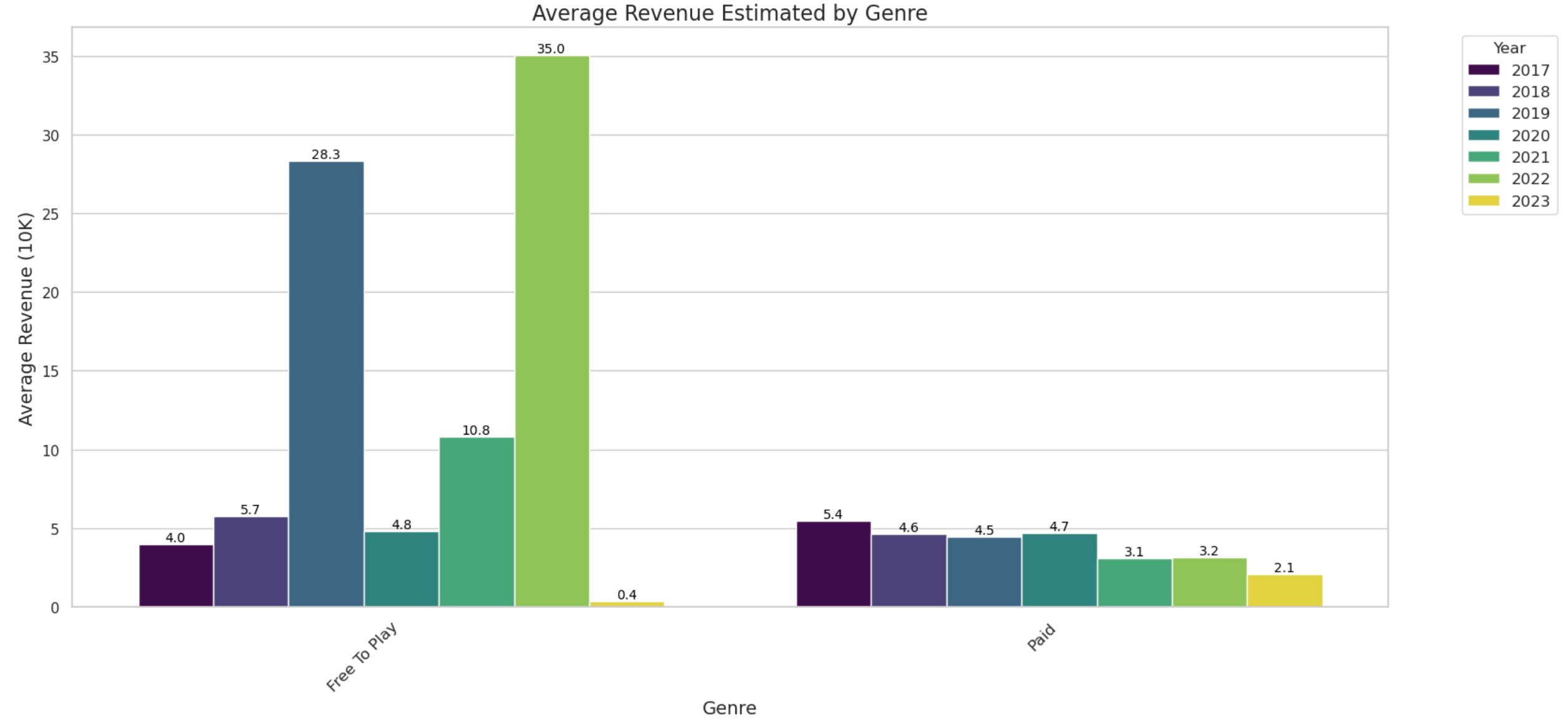
1. **Top Performers**:

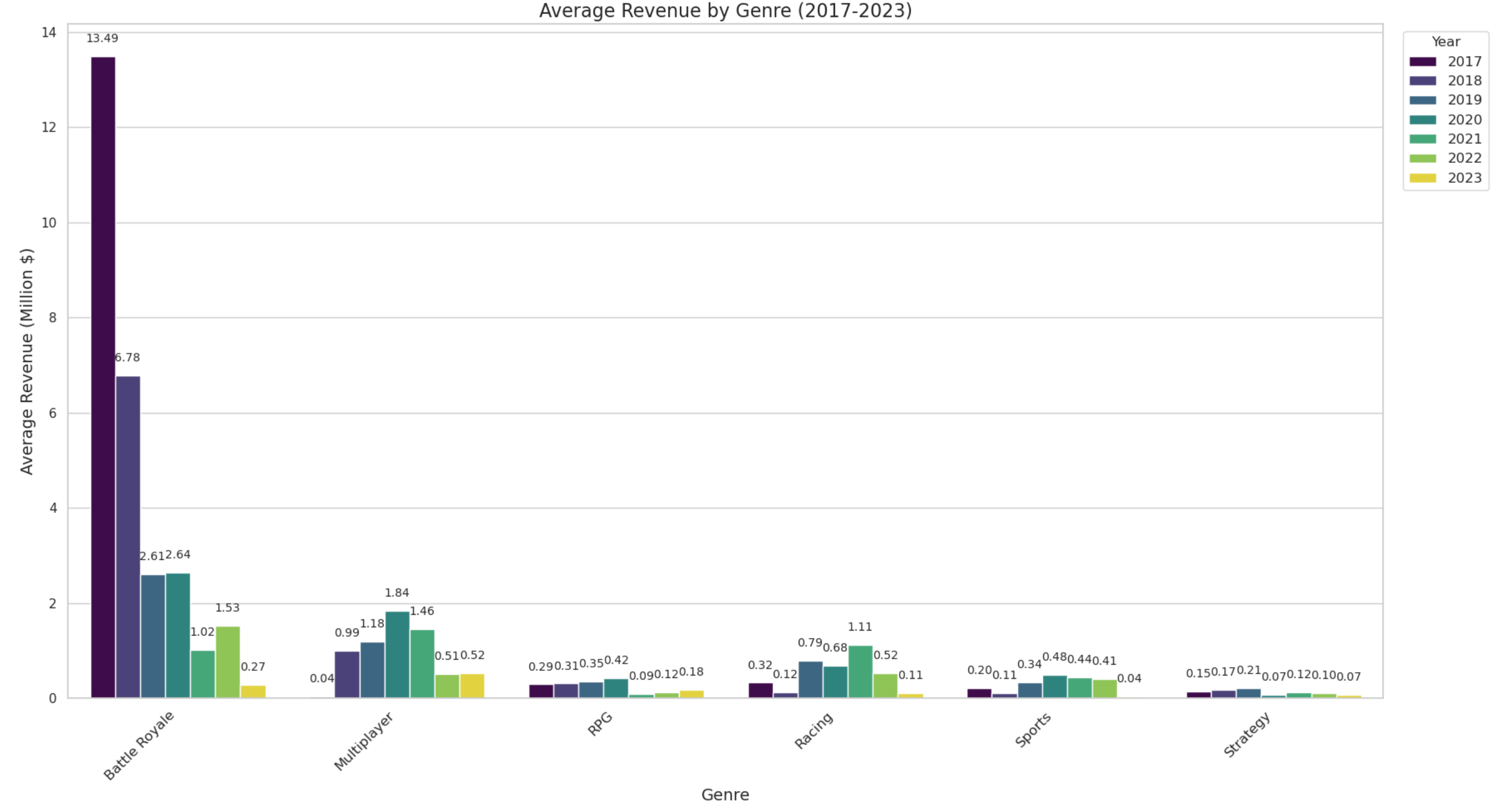
The Strategy genre has the highest average review score in the most recent year, followed closely by the Sports genre. These high scores could reflect a dedicated fan base or a smaller, more curated set of game offerings that tend to receive better reviews.

1. **Stability in Some Genres**:

While the overall trend is upwards, some genres such as RPG and Multiplayer show relatively stable review scores across the years, suggesting consistent game quality or a matured genre with established standards.

Revenue Estimate vs. Genre Trends





1. **Racing and Sports genres show moderate performance**:

Racing and sports games appear to have moderate revenue figures. These genres may appeal to niche audiences with specific interests, which could explain their middle-range performance in revenue compared to RPGs and Strategy games.

1. **Variability in Free to Play (F2P) revenue**:

The Free to Play genre shows fluctuations in revenue, which could be due to the varying success of monetization strategies in these games. Some F2P titles may rely on microtransactions or in-game advertising, which can lead to inconsistent revenue streams based on player engagement and spending behaviors.

1. **Overall trend of increasing revenue**:

A general upward trend in revenue across most genres suggests that the gaming industry is growing. It highlights the potential for increased investment in game development, marketing, and customer engagement to capitalize on this expanding market.

1. **The impact of top titles**:

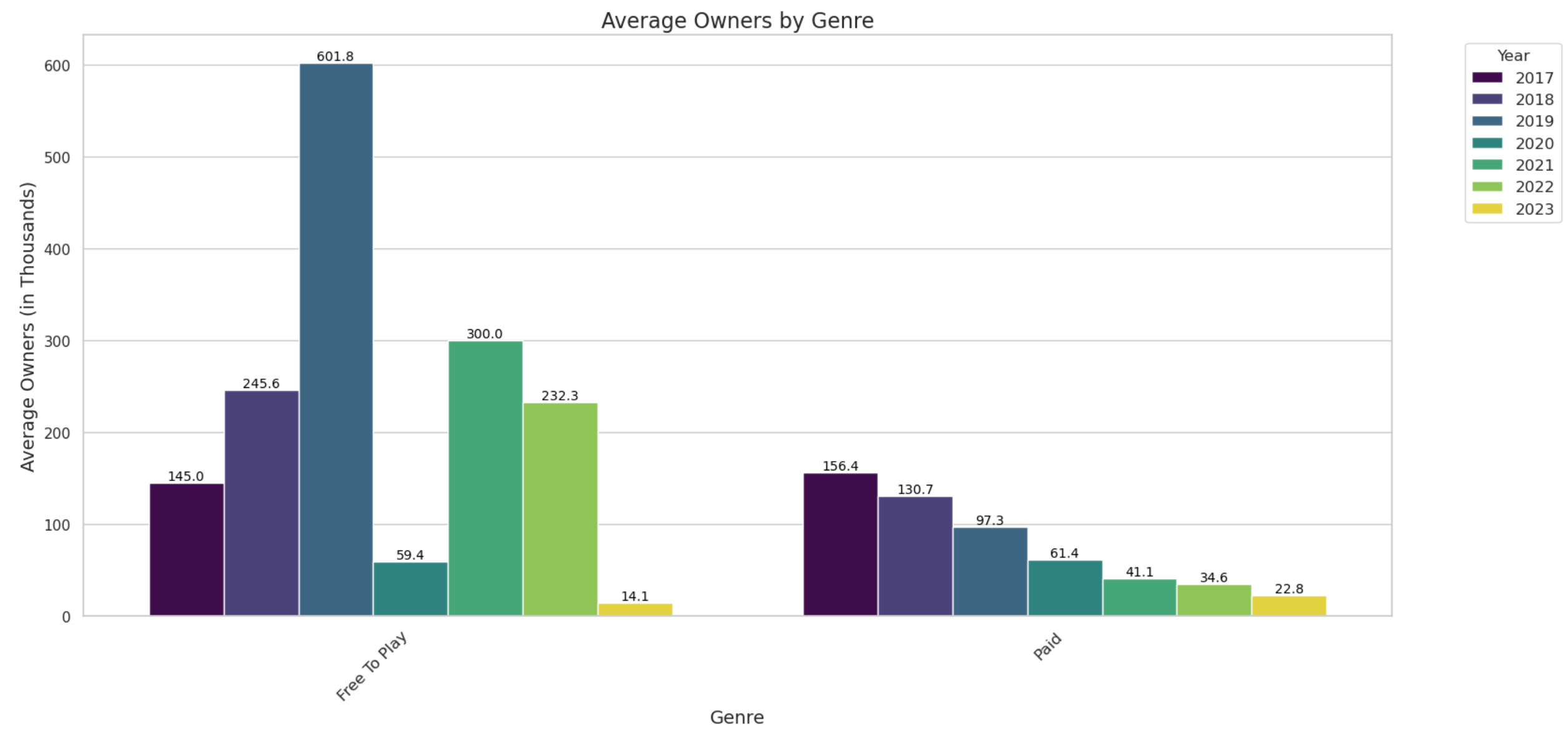
The presence of top games in each genre indicates that standout titles can significantly impact overall revenue. It may be beneficial for developers and publishers to focus on creating high-quality titles that have the potential to become hits in their respective genres.

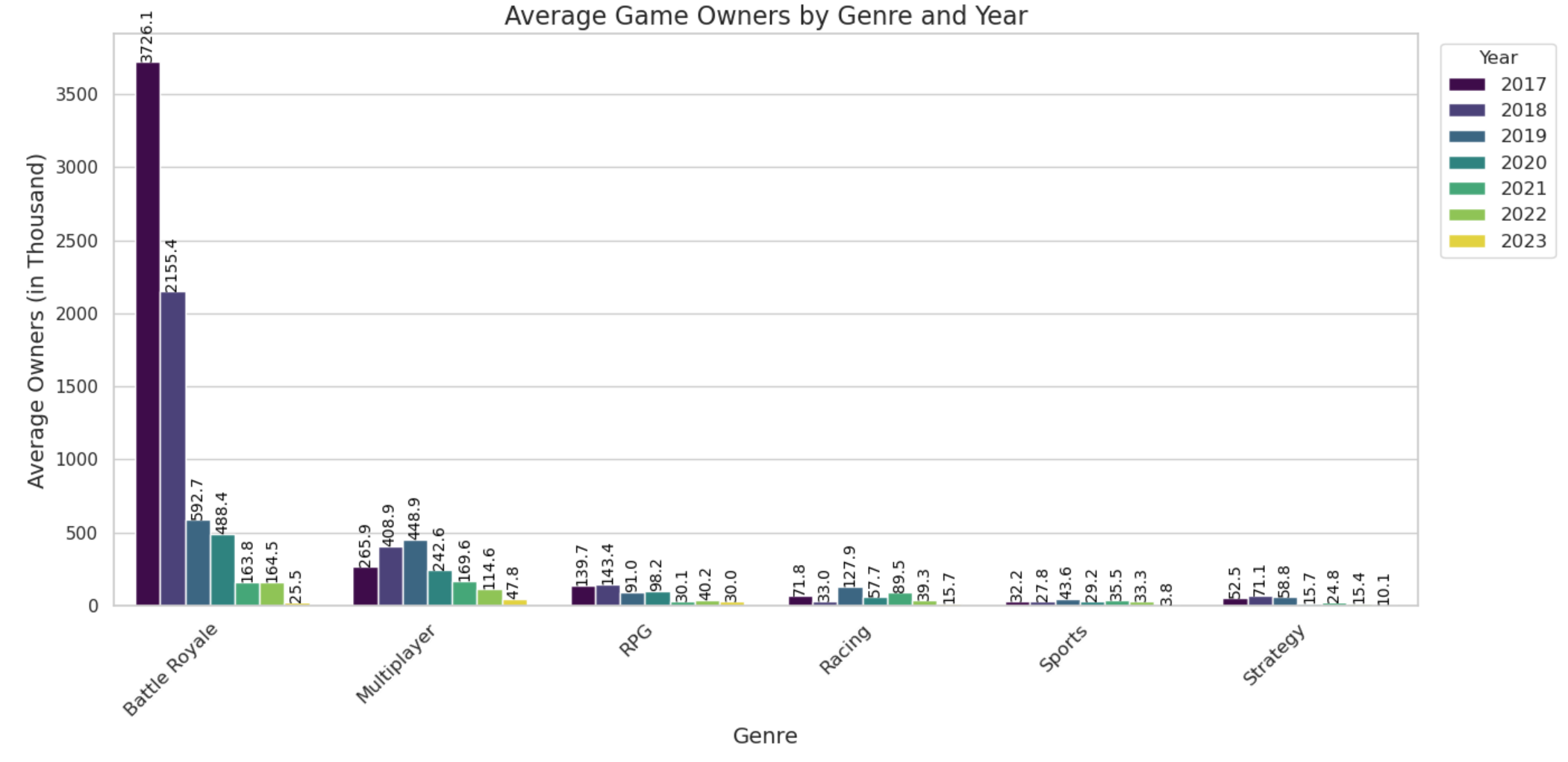
1. **Diversity in profitable genres**:

The variety among top-performing genres indicates that there's no single formula for success when it comes to genre. A diversified portfolio could be beneficial for developers and publishers to cater to different audience segments and mitigate market risks.



Ownership vs. Genre Trends





1. **Popularity Peaks and Variances**:

There is a standout peak in popularity for certain games within genres, suggesting that while average ownership might provide one picture of genre performance, individual game titles can significantly outperform the average, possibly due to virality, critical acclaim, or other factors that drive player engagement.

1. **Genre Life Cycle Trends**:

Fluctuations in the average number of game owners across years within genres may indicate the life cycle of popularity for certain game types. Some genres experience spikes in ownership that could correspond with the release of highly anticipated titles or innovations within the genre, while others may see a decline, which could be due to market saturation or a shift in player preferences.

1. **Impact of Business Models**:

There's a notable difference between the performance of Free to Play games versus Paid games, with Free to Play games generally attracting more owners. This reflects the accessibility of the free-to-play model and its effectiveness in drawing in a larger player base, which may not always translate into better review scores or sustained popularity.

1. **Evolution of Player Preferences**:

The trends also reflect an evolution in player preferences over time. Genres that once dominated in ownership numbers may see a shift as new genres gain popularity. This could be a result of technological advancements, changing demographics, or the rise of new gaming platforms that better support certain genres.



**Exploring Game Data With TF-IDF and Cosine Similarity**

Functional Blocks

**Implications & Strategic Recommendations**

Implications

1. **Market Insights for Developers and Publishers:**

* Our analysis reveals valuable insights into the shifting landscape of the gaming industry, providing developers and publishers with actionable data to inform their strategic decisions.
* Understanding genre preferences, revenue trends, and user engagement patterns can help stakeholders tailor their game development strategies to meet evolving consumer demands.

1. **Impact on Marketing and Promotion Strategies:**

* The analysis of review score trends and revenue performance offer insights into the effectiveness of marketing campaigns and promotional efforts.
* By identifying genres with higher review scores and revenue potential, marketers can allocate resources more effectively to maximize ROI and enhance brand visibility.

1. **Platform Optimization Opportunities:**

* The exploration of ownership trends across different genres sheds light on the platforms and distribution channels favored by gamers.
* Platform owners can leverage this information to optimize their services, tailor content offerings, and enhance user experience to attract and retain a larger audience.

1. **Emerging Trends in Gameplay and Monetization Models:**

* The observed variability in revenue trends among different genres highlights the importance of exploring innovative gameplay mechanics and monetization models.
* Developers can capitalize on emerging trends such as eSports integration, cross-platform playability, and in-game purchases to drive revenue growth and maintain player engagement.

#### Strategic Recommendations

1. **Genre Diversification Strategy:**

* Based on the analysis, we recommend a diversified approach to game development, balancing investments across multiple genres to mitigate risks and capitalize on emerging market opportunities.
* Developers should prioritize genres with strong revenue potential and positive review scores, while also exploring niche markets and innovative gameplay concepts to differentiate their offerings.

1. **Quality Enhancement Initiatives:**

* To meet the rising expectations of players, developers should focus on enhancing the quality and polish of their games, prioritizing factors such as graphics, gameplay mechanics, and storytelling.
* Continuous user feedback and iteration cycles should be integrated into the development process to address issues and refine features based on player preferences.

1. **Monetization Strategy Optimization:**

* Developers should adopt a strategic approach to monetization, balancing the implementation of in-game purchases and premium content with the need to maintain player satisfaction and retention.
* Leveraging data analytics and player behavior insights, developers can fine-tune their monetization strategies to offer value to players while maximizing revenue generation.

1. **Investment in Multiplayer and Competitive Gaming:**

* Given the popularity of multiplayer and competitive gaming genres, we recommend increased investment in developing titles tailored for eSports and competitive gaming circuits.
* Integrating features such as robust matchmaking systems, spectator modes, and tournament support can enhance the competitive experience and attract a dedicated player base.

1. **Cross-Platform Compatibility and Community Building:**

* To maximize audience, reach and player engagement, developers should prioritize cross-platform compatibility, allowing players to seamlessly transition between different devices and platforms.
* Community-building initiatives, such as forums, social media engagement, and live events, can foster a sense of belonging among players and enhance the longevity of game communities.
* By implementing these strategic recommendations based on the insights derived from our analysis, stakeholders in the gaming industry can position themselves for success in an ever evolving and competitive market landscape.

**Future Outlook and Enhancements**

Future Outlook

* **Sustained Interest in Free-to-Play Models**

The enduring popularity of free-to-play games is expected to persist, driven by their ability to attract a broad player base without the barrier of upfront costs. This model will likely continue to dominate the gaming landscape, especially in markets where accessibility and affordability are key factors for player engagement.

* **Continued Evolution of Battle Royale and Multiplayer Genres**

The rising trend in ownership and review scores for Battle Royale and Multiplayer games indicates their ongoing popularity. As these genres evolve, there is significant potential for innovative titles to capture a larger market share and shape the competitive gaming landscape.

* **Growth Potential for RPGs**

While RPGs have experienced fluctuations in popularity, there remains substantial room for growth. Developers can explore new immersive experiences and cross-genre innovations to rejuvenate interest in this genre and attract a diverse player base.

* **Increased Quality in Paid Games**

The trend toward higher review scores for paid games is likely to continue as developers prioritize delivering higher-quality experiences to justify the initial purchase. This emphasis on quality is expected to enhance player satisfaction and drive continued growth in the paid gaming segment.

* **Varied Success in Sports and Racing Genres**

The varied popularity of sports and racing games underscores the need for innovation and adaptation within these genres. Leveraging new technologies such as virtual reality (VR) and exploring innovative gameplay mechanics can help revitalize player interest and sustain engagement over time.

* **Strategic Opportunities for Strategy Games**

Strategy games continue to maintain a solid presence in the gaming market. Developers have strategic opportunities to capitalize on this steady interest by exploring new platforms, audiences, and gameplay mechanics to broaden their appeal and attract new players.

#### **Market Dynamics Influencing Genre Popularity**

The dynamic nature of game ownership across genres highlights the importance of adaptability in marketing and development strategies. To remain competitive, developers and publishers must closely monitor player interests and market trends, adjusting their approaches to align with evolving consumer preferences.

### Future Project Enhancements

### **Cross-Platform Performance Comparison**

Comparing genre popularity across different gaming platforms can provide valuable insights into platform-specific trends and preferences. This analysis will help developers optimize their cross-platform strategies and tailor their offerings to maximize reach and engagement.

* **Technological Advancements Influence**

Assessing the impact of emerging technologies such as augmented reality (AR), virtual reality (VR), and cloud gaming on game genre trends is crucial for staying ahead of industry shifts. Understanding how these advancements influence player preferences and behavior will inform strategic decisions and innovation in game development.

* **User Retention Metrics**

Evaluating user retention rates across different genres over time offers valuable insights into long-term engagement patterns. By identifying which genres have higher retention rates, developers can refine their content, features, and monetization strategies to foster sustained player engagement and loyalty.

* **Content Creator and Influencer Impact**

Tracking and quantifying the impact of content creators and influencers on game genre popularity provides valuable marketing insights. Collaborating with influencers and content creators can amplify brand visibility, drive player acquisition, and shape genre trends through targeted promotional campaigns and partnerships.

* **Sentiment Analysis of User Reviews**

Implementing sentiment analysis on user-generated reviews enables developers to extract deeper insights into player sentiments, preferences, and concerns regarding different game genres. This qualitative data can inform product improvements, content updates, and community engagement strategies to enhance player satisfaction and retention.

* **Gameplay Duration Metrics**

Collecting and analyzing data on total gameplay duration offers an alternative perspective on game trends and player engagement. Understanding which genres keep players engaged for longer periods provides valuable feedback for developers seeking to optimize gameplay experiences, improve retention, and maximize player enjoyment.

**Conclusion**

This project aims to provide actionable insights into the dynamics of video game genres, leveraging data analytics techniques to uncover trends, patterns, and factors influencing genre preferences among gamers. By addressing the outlined objectives and goals, we anticipate contributing valuable knowledge to industry stakeholders, facilitating informed decision-making, and driving positive outcomes in the ever-evolving gaming landscape.

**References**

1. VG Insights

https://vginsights.com/insights/article/how-to-estimate-steam-video-game-sales

1. Newsletter

https://newsletter.gamediscover.co/p/how-that-game-sold-on-steam-using?s=r

1. Steam
2. Kaggle

**Appendix**

Technologies & Tools Used:

1. PySpark
2. Amazon Web Services (AWS)
3. Spark-SQL
4. JupyterHub/ JupyterNotebook
5. Docker
6. ChatGPT 3.5 (Editing the Report and Slides)

* Utilized ChatGPT to mostly edit our report and slide content. We also used it to perform minor code debugging, learning important concepts for creating our project.