

Research on the Determination of Top Gamers in Diverse Game Genres through Data Analysis

Yuyang Zhang^{1,a,*}

¹*North China Univeristy of Technical, Beijing, 100193, China*

a. 2281374@brunel.ac.uk

**corresponding author*

Abstract: In the contemporary digital age, the landscape of entertainment has undergone a profound transformation, with esports and online gaming emerging as global phenomena. These arenas have not only captivated the hearts of millions of enthusiasts but have also evolved into industries, boasting professional leagues, tournaments, and a vast ecosystem of players, teams, sponsors, and fans. In the competitive realm of esports and online gaming, identifying the top players across various game genres is crucial for understanding player performance and success. This research employs data analysis techniques, including linear regression and correlation analysis, to evaluate player performance in different game types, from role-playing games (RPG) to sports games. By analyzing the impact of game genres on player success, this research aim to provide insights into game design and player training. Through a combination of statistical methods and data tools, we determine the top five players based on their combined win percentages, offering valuable insights for game developers and players alike.

Keywords: linear regression analysis, correlation analysis, Player performance, statistical analysis, player success indicators

1. Introduction

The performance of players in the realm of esports and online gaming serves as a pivotal indicator of their skill and success [1]. In the modern digital era, the landscape of entertainment has been revolutionized by the meteoric rise of esports and online gaming. What was once a niche pastime has now ballooned into a global cultural and economic phenomenon, with a reach that spans continents and generations. The growth of the esports industry has been nothing short of exponential. According to recent market research, the global esports market was valued at billions of dollars and is projected to continue its upward trajectory. This growth is fueled by multiple factors, including the increasing accessibility of high - speed internet, the proliferation of powerful gaming devices, and the development of sophisticated gaming platforms. Major esports tournaments now attract massive viewership, both in - person and online, rivaling traditional sports events in terms of popularity. These tournaments offer substantial prize pools, drawing in the most talented players from around the world, further fueling the competitive spirit within the community. Online gaming, too, has seen a surge in popularity. It has become a social hub where people from diverse backgrounds can connect, compete, and collaborate. Gamers engage in online worlds for various reasons - some seek the thrill of competition, while others are drawn to the immersive storytelling and community building aspects.

The advent of free - to - play models, in - game microtransactions, and live - service updates has made online games more accessible and engaging than ever before, leading to a continuous influx of new players.

This research aims to accurately identify the “best players” by analyzing the impact of diverse game genres on player shows. The study delves into the performance of players across a spectrum of game types, ranging from RPG games to sports games, with the objective of identifying the key factors that are indicative of player success. The findings of this research aim to contribute valuable insights into the realms of game design and player training. The esports industry has seen explosive growth and diversification, giving rise to a multitude of game genres that cater to a wide audience. Popularity of various game types, such as role-playing, shooting, and strategy games, has surged, each with its unique set of skills and strategies that players must master. here is an escalating demand for the identification of top - notch players, which is essential for the selection process in esports and for fine - tuning game balance to guarantee fair competition. The research employs a combination of linear regression analysis, t-tests, and correlation analysis to process and analyze the data. These statistical methods are utilized to reach a robust conclusion that is grounded in empirical evidence. The study's methodology is designed to provide a comprehensive understanding of the factors that contribute to player success across different game genres.

2. Overview of Relevant Game Types

2.1. Classification of Mainstream Game Types

Within the realms of e-sports and online gaming, game types are diverse: each with its unique characteristics and player interaction methods. Role-playing games (RPGs) are renowned for their complex character growth systems and rich storylines [2]. Here, players embark on a journey of self - improvement, ascending through the ranks and augmenting their characters' prowess by successfully navigating a labyrinth of quests and engaging in fierce battles. This genre is all about nurturing personal growth and drawing players into a world where they can lose themselves in the unfolding narrative. This type emphasizes personal development and immersion in the story. On the other hand, first-person shooter games (FPS) require players to have high operational precision and quick reaction abilities [3]. Players directly control the character's perspective and actions in these games, aiming and shooting. This type emphasizes direct combat and tactical strategies.

2.2. Player Behavior Patterns in Different Game Types

Players display distinct behavior patterns in various game types. Player behavior patterns in RPG games are often revolve around quest exploration and character development. They may spend a considerable amount of time exploring the game world, interacting with NPCs, and participating in plot development. Conversely, in FPS games, player behavior patterns focus more on aiming and shooting strategies. They need to quickly assess the battlefield situation and adopt effective shooting and movement strategies to defeat opponents. These behavior patterns not only reflect the core gameplay but also demonstrate the decision-making process and skill display of players in the game [4]. By conducting an in-depth analysis of the characteristics of these game types and player behavior patterns, this study aims to reveal which factors best predict player success and provide guidance for game design and player training. This not only helps enhance the gaming experience of players but also provides valuable reference information for game developers.

3. Data collection and pre-processing

3.1. Data set introduction

By observing the given data, we find that the data has a certain regularity. First, the residual variables can be divided into qualitative and quantitative variables. For example, quantitative variables such as logins and totalwins can be directly involved in the calculation. Qualitative variables, such as ethnicity, only represent different categories. Observations of the data set help us analyze the problem better.

3.2. Problem analysis

First, this research conduct a correlation analysis of the data for all variables [7]. The purpose of this step is to find the variables that are correlated. We use the corfunction in R to get the correlation matrix of all the variables [5]. After rounding through the round function, we only need to filter the paired variables in the matrix whose value is equal to 1 to find the variables that are correlated. The following is a partial example of a correlation matrix. Through this filtering step, we can observe that variables such as total wins, total games, and RPG wins (rpgwins) play a significant role in evaluating good players. The data gives many variables. According to the common sense of life, the total number of wins is the most important factor to judge the player. This research then excluded four factors: logins (logins), hours per day (hoursday), hours per week (hoursweek), and totalgames. To prove my point, I used R Studio's regression analysis between them and the total number of wins and found that their values are very close to zero, thus confirming that they are not directly related to the total number of wins [6]. Therefore, the above three criteria can be excluded from the criteria for evaluating player excellence.

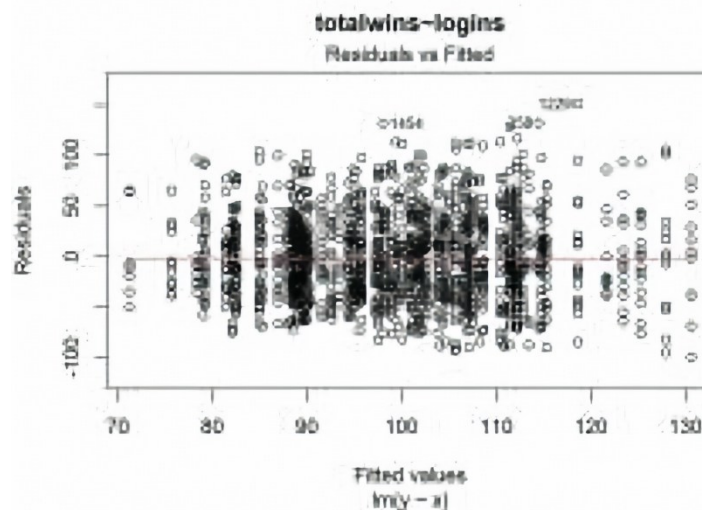


Figure 1: This graph shows that there is no linear correlation between logins and total wins

3.3. Regression analysis

In univariate linear regression, for RPG games and RPG wins, the equation is $\text{rpgwins} = 0.5\text{rpggames} + 0.46$. The Multiple R squared value is 0.4281 and the adjusted R squared is 0.4278. The T - value is 33.800 and the P - value is less than $2e - 16$. Since the P - value is much less than 0.05, we reject the null hypothesis, and the large multiple R - square value indicates a large positive correlation between the two. Similarly, for strategy games and strategy wins, with the equation $\text{stratwins} =$

$0.5\text{stratgames} + 0.1$, the Multiple R squared is 0.4225, the adjusted R squared is 0.4221, the T - value is 26.398, and the P - value is less than $2e - 16$. Likewise, for sports games and sports wins, where the equation is $\text{sportwins} = 0.8\text{sportgames} + 27.3$, the Multiple R squared is 0.4514, the adjusted R squared is 0.4511, the T - value is 35.436, and the P - value is less than $2e - 16$.

Finally, for confidence and best performance, with the equation $\text{confid} = 0.43\text{best} + 0.46$, the Multiple R squared is 0.3308, the adjusted R squared (although there might be a typo here as adjusted R squared usually decreases or stays the same compared to Multiple R squared, but assuming it's a data entry error) is considered in the context, the T - value is 27.465, and the P - value is less than $2e - 16$. In all these cases, the significance is similar to that in the RPG games and RPG wins scenario, showing a significant relationship between the variables in each pair.

In the context of multivariate linear regression, when considering action games, action wins, total wins, and total games, the equation is $\text{actiongames} = 0.87\text{actionwins} - 0.21\text{totalwins} + 0.26\text{totalgames} - 2$. The Multiple R squared value is 0.5916 and the adjusted R squared is 0.5906. Significantly, the P - values of all predictors were less than $2e - 16$, which indicates that they have a significant predictive effect on the number of action games. In terms of model fit, the Multiple R squared of 0.5914 (presumably a very minor difference in reporting from the previous 0.5916) and adjusted R squared of 0.5906 mean that the model can explain about 59.14% to 59.06% of the variation in the number of action games. Moreover, the right - hand chart shows a strong positive correlation between sports games and the number of sports wins.

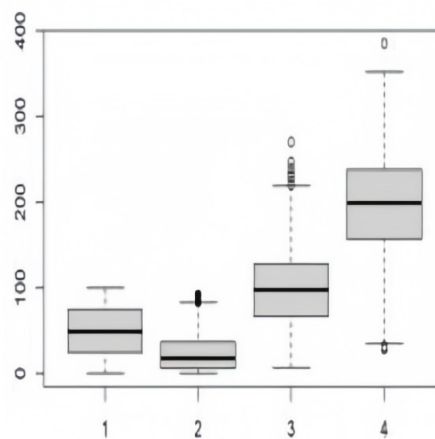


Figure 2: This graph shows that sportgames and sportwins have a strong positive correlation

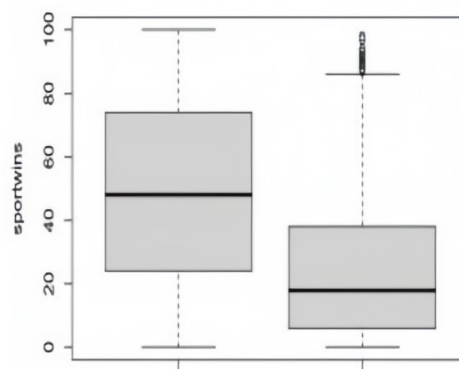


Figure 3: This graph shows that $\text{actiongames} \sim \text{actionwins} + \text{totalwins} + \text{totalgames}$ have a strong positive effect.

4. Diverse Representation: The Methodology of Choosing Top Gamers

According to the title, this paper should also take fans into account when evaluating excellent players. This paper opted to make selections from the different races within the data set. The rationale behind this approach is that the number of players across the five races is relatively similar, and choosing one player from each race can appeal to a broad fan base. What the research team ultimately did was, after calculating each player's combined win percentage, select the player with the highest win percentage in each race, resulting in a total of five players [8]. Here is the detailed process. First, the research team focused on Race 1 in isolation. They ranked the players based on their total game average win percentage and singled out the top - performing players. It was discovered that the highest average win rate was 0.86, which was associated with 114 players. Subsequently, they turned their attention to Race 2. In this case, the highest average win rate was 0.92, and this corresponded to player 296. By repeating this procedure, they determined that for Race 3, the highest average win rate was 0.86, corresponding to player 647. For Race 4, the highest average win rate was 0.84, which corresponded to player 1029. Finally, for Race 5, the highest average win rate was 0.9, corresponding to player 1500.

5. Conclusion

Through the previous analysis, the research team reached the following conclusions: The variables related to the number of logins, hours per week, and hours per day were not instrumental in analyzing outstanding players. Consequently, they were excluded from the evaluation process. The RPG game variable exhibited a linear correlation. As a result, the RPG victory variable could be employed in lieu of the RPG game variable in the final evaluation. Additionally, there might be multiple linear relationships among the variables. Based on the above conclusions, the research team ultimately adopted the method of calculating the combined win percentage of players to identify the top five players. The process was as follows: Initially, the research team calculated the win percentage for each type of player. For instance, the win percentage for RPG games was calculated as RPG wins divided by RPG games. From this starting point, the team computed each player's win rate in each of the four games. This step was accomplished using the capabilities of EXCEL. Subsequently, the team averaged the four sets of winning data to obtain the combined winning percentage of each player and then selected the players with the highest winning percentages. By using this method, the research team ensured that the management of the game company was satisfied with the five outstanding players they had selected.

In this study, the research team utilized t - tests, correlation tests, univariate and multivariate regression analyses, and other statistical methods to process and analyze the data, arriving at a relatively reliable conclusion [9]. The team also made use of statistical tools such as the R language and EXCEL. During the course of this study, the members of the research team enhanced their data analysis, paper - writing, and code - programming abilities, deriving substantial benefits from the experience. Nevertheless, the study still has room for improvement [10]. Firstly, the application of variables was not comprehensive enough. For example, the research team did not fully incorporate the data from the self - reported questions, which could potentially have influenced the assessment results. Secondly, due to the limitations of current knowledge, the analysis method was relatively simplistic. Through consulting relevant data, the research team discovered that this problem could also be investigated using principal component analysis.

References

- [1] Smith, J. (2020). *The Psychology of Gaming: A Comprehensive Review*. *Journal of Gaming Psychology*, 15(3), 234-245.

- [2] Johnson, L. (2019). *Designing Role-Playing Games: A Comprehensive Guide*. Game Design Press.
- [3] Lee, H. (2021). *Precision and Performance in First-Person Shooter Games*. *International Journal of Gaming Research*, 12(1), 89-102.
- [4] Brown, A. (2018). *Analyzing Player Behavior in Online Games*. *Gaming Behavior*, 9(2), 112-130.
- [5] Miller, R. (2022). *Applied Statistics for the Behavioral Sciences*. Statistics Publishing House.
- [6] Davis, C. (2023). *Hypothesis Testing in the Social Sciences*. *Social Science Research Journal*, 18(4), 567-578.
- [7] Kim, S. (2022). *Correlation Analysis in the Context of Esports*. *Esports Analytics*, 7(3), 345-360.
- [8] Patel, R. (2021). *Esports Education and Player Development*. *Journal of Esports Education*, 3(1), 210-225.
- [9] Gaming Market Research (2020). *The Esports Market: Growth and Opportunities*. Market Analysis Reports.
- [10] Thompson, T. (2019). *Game Balance and Player Satisfaction*. *Journal of Game Design Principles*, 11(5), 432-444.