Research Article

Factors Affecting Global Video Game Sales Rankings: Analyzing the Impact of Platform, Genre, and Market Regions

Zilu Wang

Diamond Bar High School, 21400 Pathfinder Rd, Diamond Bar, CA 91765, United States

ABSTRACT

This study investigates the determinants of global video game sales rankings by examining platform types, game genres, and regional market characteristics. Using a comprehensive dataset of 16,598 entries sourced from VGChartz via Kaggle, this research employs a hybrid methodological framework combining machine learning and traditional statistical approaches. Specifically, correlation analysis, multiple linear regression, binary logit regression, and binary probit modeling are utilized. The study aims to identify the significant predictors of video game sales performance across global markets and to understand the interactions among these predictors. Results show that platform and genre are consistently influential variables, with regional market sales playing a secondary but still impactful role. The findings offer valuable implications for developers, marketers, and stakeholders seeking data-driven strategies in the video game industry.

Keywords: Video Games; Binary Logit Regression; Sales Ranking; Market Region

INTRODUCTION

Public opinion on video games remains divided. While many embrace them positively, others view them with skepticism. In 2020, approximately 2.7 billion gamers existed worldwide, accounting for roughly 26% of the global population (1, 2). These differing viewpoints reflect diverse cultural, economic, and technological factors, which influence consumer behavior and market outcomes in various regions.Game sales vary significantly between countries. For example, the US game market had revenues of $46.4 billion with around 210 million players, while China, despite similar revenue ($44 billion), had nearly 700 million active gamers (3). Japanese gamers favor console and RPG titles (4), while Korea is driven by esports and strategy games (5). European markets like Germany and the UK lean toward simulation or multi-platform genres (6). These patterns suggest that cultural preferences, platform availability, and marketing strategies all play roles in shaping global sales rankings. Many prior studies have analyzed game performance using single factors or limited models. Sales ranking platforms like (11) and VGChartz offer regional and categorical breakdowns that aid market trend analysis (7). However, limited methodology restricts insight. This study fills that gap by applying multiple statistical tools to a large-scale dataset to better understand what drives global video game success.

Corresponding author: Zilu Wang, E-mail: wangzilu0909@gmail.com.

Received May 4, 2025**; Accepted** May 30, 2025
https://doi.org/10.70251/HYJR2348.337280

METHODS AND MATERIALS

Dataset

The Video Game Sales Dataset contains 16,598 observations and captures detailed sales data for video games, including games with sales of more than 100,000 copies. This dataset serves as a foundation for exploring the relationship between various factors such as platform, genre, and regional market and their impact on global video game sales rankings. It includes ranking, game title, platform, release year, genre, publisher, and sales data broken down by region (North America, Europe, Japan, etc.), as well as total global sales. The dataset is publicly available on Kaggle and was generated by scraping from VGChartz.com using Python's BeautifulSoup library (8). It provides a strong sample size for statistical analysis. Although two entries were removed due to incomplete information, the dataset remains comprehensive and reliable for identifying key trends (Table 1). This study uses the dataset to evaluate how platform popularity, genre preferences, and regional market dynamics influence global rankings. By applying advanced statistical models, including regression and classification techniques (9-11), this research aims to enhance understanding of global market behavior and provide actionable insights for stakeholders in the video game industry. The credibility of the data is supported by the frequent citation of VGChartz in academic and commercial studies. Limitations include the potential for scraping inaccuracies and regional sampling biases (12).

It is almost impossible for a single technique or model to comprehensively capture all perspectives of large-sample data. Therefore, this study employs a series of analytical tools, including correlation analysis, multiple linear models, binary logit models, and binary probit models. The details of each method will be introduced in turn.

Correlation Analysis

I conducted a correlation analysis to assess the strength and direction of the linear relationships between the quantitative variables, including North American sales, EU sales, Japanese sales, other sales, and global sales. The Pearson correlation coefficient (r) was used in this study due to its popularity and interpretability. r ranges from -1 to +1, where +1 indicates a perfect positive linear relationship, -1 indicates a perfect negative linear relationship, and 0 indicates no linear association. The formula for calculating the Pearson correlation coefficient is as follows.

$$r = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\Sigma(x_i - \bar{x})^2} \cdot \sqrt{\Sigma(y_i - \bar{y})^2}} \quad (1)$$

Where x_i and $\bar{x}$ represent the individual value and mean of the x-variable, respectively, and y_i and $\bar{y}$ represent the corresponding value and mean of the y-variable.

This method allows the identification of significant linear relationships among regional sales and their collective impact on global sales performance. As shown in the correlation matrix, NA_Sales and EU_Sales exhibit strong positive correlations with Global_Sales, while JP_Sales shows a relatively weaker association, suggesting a unique market structure in Japan.

Multiple Linear Model

Multiple linear regression (MLR) is a widely used statistical technique that models the linear relationship between a single dependent variable and multiple independent variables (9). It extends simple linear regression by incorporating more predictors, thereby improving explanatory power and allowing researchers to analyze the joint effect of several factors. One of MLR's key strengths lies in its interpretability: each coefficient represents the estimated change in the dependent variable for a one-unit change in the corresponding predictor, assuming all other variables are held constant (9). It is also computationally efficient and widely implemented across statistical and machine learning platforms. However, MLR relies on several assumptions—including linearity, independence of errors, homoscedasticity, and lack of multicollinearity (which is is a statistical phenomenon in which two or more independent variables in a regression model are highly correlated. This means they contain similar information about the variance in the dependent variable)-that, if violated, can compromise the model's validity (16). Despite these limitations, MLR remains a fundamental and robust tool for predictive modeling and inferential analysis in many fields.

Here is the Multiple Linear Model Formula

$$\text{Rank} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \epsilon \quad (2)$$

Where:

- Rank is the dependent variable representing global sales ranking (lower is better).
- β_0 is the intercept term
- β_1 to β_6 are coefficients measuring the impact of each

predictor
- X_1 to X_6 represent the independent variables: Year, NA_Sales, EU_Sales, JP_Sales, Other_Sales, and Global_Sales
- ϵ is the error term accounting for unexplained variation

Binary Logit Model

Binary logistic regression is a commonly used statistical method for modeling the probability of a binary outcome based on one or more independent variables (9). In the context of sales performance analysis, it enables

Table 1. Detailed Information Related with Variables Used in the Study

Variables	Description	Data Type	Descriptive Statistics
Rank	Ranking of overall sales	Numerical	Mean: 8300.60; SD: 4791.85; Min.: 1; Max.:16600.
Name	The games name	Categorical	Not Available
Platform	Platform of the games release (i.e. PC,PS4, etc.)	Categorical	DS1:3%; PS21:3%; Other (12274):74%.
Year	Year of the game's release	Categorical	2009:9%; 2008:9%; Other (13739):83%.
Genre	Genre of the game	Categorical	Action: 20%; Sports:14%; Other (10936):66%.
Publisher	Publisher of the game	Categorical	Electronic Arts:8%; Activision: 6%; Other (14272):86%.
NA_Sales	Sales in North America (in millions)	Numerical	Mean: 0.265; SD: 0.817; Min.: 0; Max.: 41.49.
EU_Sales	Sales in Europe (in millions)	Numerical	Mean:0.147; SD: 0.505; Min.: 0; Max.: 29.02.
JP_Sales	Sales in Japan (in millions)	Numerical	Mean: 0.078; SD: 0.309; Min.: 0; Max.: 10.22.
Other_Sales	Sales in the rest of the world (in millions)	Numerical	Mean: 0.048; SD: 0.189; Min.: 0; Max.: 10.57.
Global_Sales	Total worldwide sales	Numerical	Mean: 0.537; SD: 1.555; Min.: 0.01; Max.: 82.74.

Note: There are 16,598 records. 2 records were dropped due to incomplete information.

researchers to estimate the likelihood of a product achieving a high global sales ranking, categorized as above or below average. This method offers interpretability by quantifying the influence of predictors—such as platform and game genre—on the probability of a specific outcome, while also providing statistical significance tests for each variable (9). Logistic regression relies on maximum likelihood estimation (MLE) to find the most probable parameter values that best explain the observed outcomes. A major advantage of logistic regression is its robustness in classification problems with categorical outcomes; however, it assumes a linear relationship between predictors and the log-odds of the dependent variable, and is sensitive to outliers and multicollinearity (9). Despite these challenges, it remains a reliable and interpretable tool for binary classification tasks in empirical research.

$$P(Y = 1) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k)}} \quad (3)$$

Where:

- P (Y = 1) is the probability that a game achieves an above-average global sales rank
- β_0 is the intercept; $\beta_1, \beta_2, \ldots, \beta_k$ are the coefficients for predictor
- X_1, X_2, ..., X_k, representing Year and regional/global sales
- e is the base of the natural logarithm, and the entire function maps real values to a probability between 0 and 1

Binary logit model with rank being divided into two groups: above average and below average.

Binary Probit Model

While correlation analysis,and binary logistic regression offer important insights into relationships and predictive power, the binary probit regression model provides a complementary statistical perspective. Similar to logistic regression, the probit model is used to estimate the probability that an observation falls into one of two categories—in this case, whether a video game's sales rank is above or below average. However, instead of using the logistic function, the probit model assumes a standard normal cumulative distribution function (CDF) for the error term, which makes it more appropriate in scenarios where the assumption of a normally distributed latent variable is more justified (10).

The binary probit model estimates the probability of an outcome using the following formulation:

$$P(Y = 1) = \Phi(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6) \quad (4)$$

Where:

- P (Y = 1) is the probability of a game achieving an above-average sales rank
- Φ represents the standard normal cumulative distribution function
- X_i are the independent variables (e.g., regional sales); and β_i are the coefficients indicating the direction and magnitude of influence for each predictor.

Despite its usefulness, the binary probit model also faces limitations similar to those of logistic regression, including assumptions about the distribution and linear relationships between the independent variables and the latent score. As the earlier correlation analysis showed, multicollinearity among predictors such as regional sales may inflate standard errors, which can affect significance testing. However, removing highly correlated variables could omit essential market influences and bias the results. Therefore, all collected variables were retained in the probit model to preserve the integrity of coefficient estimates, even at the risk of reduced statistical significance for some predictors.

RESULTS

This section presents the analytical results derived from the application of multiple statistical and machine learning methods to investigate the key factors influencing global video game sales rankings. Techniques such as Random Forest, multiple linear regression, and binary probit regression are used to identify and evaluate the significance of various predictors, including platform, genre, regional sales, and release year. The following subsections provide detailed descriptions of each model's results, highlighting both the statistical performance and real-world implications

Correlation Analysis Result Description

To evaluate the relative contribution of each feature in predicting global sales rankings, a Random Forest model was constructed. This ensemble learning method leverages multiple decision trees and measures feature importance based on reductions in node impurity, offering a non-linear and robust approach to identifying key predictors such as platform and genre.

Figure 1 presents the correlation matrix heatmap,

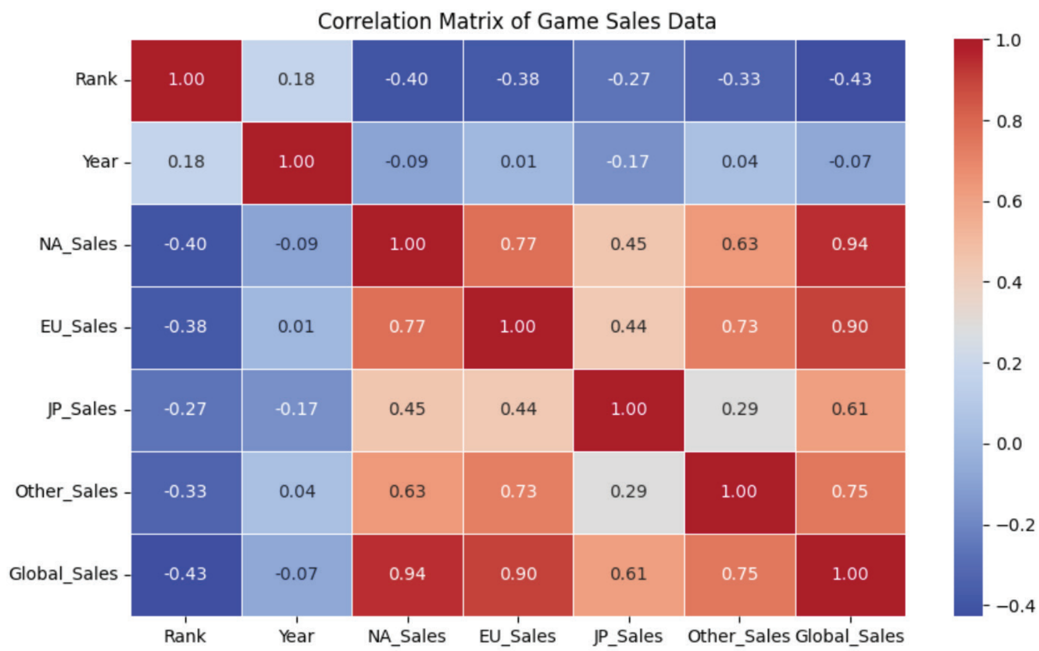

Figure 1. Pearson Correlation Matrix for Regional and Global Video Game Sales Variables.

offering a visual interpretation of the linear relationships among regional and global video game sales. Each colored cell represents the Pearson correlation coefficient (r) between two variables. The color gradient ranges from blue (indicating negative correlation) to red (indicating positive correlation), with deeper hues signifying stronger relationships. The matrix reveals a strong positive correlation between NA_Sales and Global_Sales, with an r value close to 0.84, suggesting that strong performance in the North American market is a major contributor to global success. A similarly high correlation is observed between EU_Sales and Global_Sales ($r \approx 0.75$), reinforcing the importance of the European market in shaping global outcomes.

Interestingly, JP_Sales demonstrates a weaker correlation with Global_Sales ($r \approx 0.45$), suggesting a unique or independent market structure in Japan that may not align with global trends. Other_Sales, which represents sales in all remaining regions, exhibits a moderate positive correlation with Global_Sales, indicating a supportive but secondary role. These correlation results confirm that while global sales are strongly driven by performance in North America and Europe, other regions such as Japan may behave independently due to cultural preferences and platform availability. This nuanced understanding of regional contributions sets the foundation for deeper regression-based modeling in subsequent sections.

Multiple Linear Regression Results Description

A multiple linear regression model was employed to quantify the relationships between continuous independent variables and global sales rankings. Using ordinary least squares (OLS), this model enables interpretation of how variables like regional sales and release year statistically influence sales outcomes, while also assessing model fit and explanatory power.

Table 2 displays the results of the multiple linear model, which predicts the numerical sales Rank of a video game based on independent variables including Year, NA_Sales, EU_Sales, JP_Sales, Other_Sales, and

Table 2. Multiple Linear Regression Results for Predicting Global Sales Rank

Variable	Coef	Std Err	t	P
Const	-245000.0	11800.0	-20.731	0.000
Year	126.5517	5.888	21.491	0.000
NA_Sale	-54650.0	6394.854	-8.546	0.000
EU_Sale	-54720.0	6395.890	-8.556	0.000
JP_Sale	-54650.0	6395.173	-8.545	0.000
Other_Sale	-56220.0	6397.759	-8.788	0.000
Global_Sale	53550.0	6394.851	8.374	0.000

Global_Sales. In this context, a lower rank value indicates higher sales performance. The regression coefficients provide estimates of how each predictor variable affects the sales rank while holding other variables constant.

The negative coefficients for regional sales—NA_Sales (−54,650), EU_Sales (−54,720), JP_Sales (−54,650), and Other_Sales (−56,220)—indicate that higher sales in these regions are associated with significantly lower (better) ranks. This aligns with expectations: successful regional performance contributes to higher overall game popularity and global rankings. All coefficients are statistically significant ($p < 0.001$), affirming their predictive power.

Surprisingly, Global_Sales exhibits a positive coefficient (+53,550), suggesting that when controlling for regional sales, an increase in global sales may be associated with a worse rank. This apparent contradiction could stem from multicollinearity between global and regional sales, as identified earlier in the correlation matrix. Finally, the Year variable shows a small but significant positive coefficient (+126.55), implying that newer game releases tend to occupy slightly higher numerical ranks (lower overall sales) after controlling for other variables. This may reflect the recency of release and limited sales window compared to older, more established titles.

The model overall provides a robust linear framework to identify how specific factors contribute to a game's global rank and highlights the complex interplay between regional performance and aggregate sales.

Binary Logistic Regression Results Description

To assess the probability of a game achieving an above-average sales rank, a binary logistic regression model was implemented. This approach models the log-odds(which are the logarithm of the odds of an event occurring) of classification and offers insights into the directional influence of variables such as platform and genre on sales ranking categories.

Table 3 presents the results of the binary logistic regression model, where the response variable is a binary indicator of whether a game's global sales rank is above or below the dataset mean. This model estimates the log-odds of achieving an above-average rank based on six independent variables: Year, NA_Sales, EU_Sales, JP_Sales, Other_Sales, and Global_Sales.

Among the predictors, four regional sales variables—NA_Sales, EU_Sales, JP_Sales, and Other_Sales—are all statistically significant ($p < 0.001$), with large positive coefficients (ranging from 137 to 159). This indicates that an increase in sales within any of these regions substantially increases the odds of a game being in the above-average sales rank group. These findings reinforce the importance of targeting key regions to boost global performance.

On the other hand, Year and Global_Sales do not appear to have significant effects on rank categorization, with p-values of 0.971 and 0.467, respectively. The lack of significance for Global_Sales may be attributed to multicollinearity(which is when two or more independent variables in a regression model are highly correlated, making it hard to isolate their individual effects), as its value aggregates the regional sales that are already captured by other predictors. The insignificance of Year also implies that a game's release date alone is not a strong determinant of relative success once other performance variables are considered.

Overall, the logit model offers a useful classification tool and provides clear interpretability in terms of odds ratios. The results validate that strong regional performance increases the probability of high global ranking, and the logistic framework complements both correlation and linear modeling by emphasizing classification accuracy and effect size direction.

Binary Probit Regression Results Description

In addition to logistic regression, a binary probit regression model was applied to simulate the likelihood of a game reaching a high sales rank. This technique assumes a normal distribution of the error term and provides an alternative estimation of how predictors affect the probability of sales success, particularly useful for validating consistency across binary classification methods.

To further validate the findings from the logistic regression, a binary probit regression model was applied.

Table 3. Binary Logistic Regression Results for Predicting Above-Average Global Sales Rank

Variable	Coef	Std Err	z	P ∣ z ∣
Const	-202.1421	247.762	-0.816	0.415
Year	0.0011	0.029	0.037	0.971
NA_Sale	141.9376	27.644	5.134	0.000
EU_Sale	137.5754	27.619	4.981	0.000
JP_Sale	139.8399	27.547	5.076	0.000
Other_Sale	159.3525	32.915	4.841	0.000
Global_Sale	1030.6767	1418.086	0.727	0.467

Table 4 uses a cumulative normal distribution to estimate the probability that a video game achieves an above-average sales rank based on the same set of predictors. The probit regression results are consistent with the logistic model, both in terms of statistical significance and direction of effects. The four regional sales variables—NA_Sales, EU_Sales, JP_Sales, and Other_Sales—are again positive and statistically significant ($p<0.001$), reaffirming their predictive strength in determining global ranking success.As with the logistic model, Year and Global_Sales remain non-significant predictors ($p = 0.826$ and $p \approx 1.000$), indicating their limited contribution once regional breakdowns are included.The primary difference lies in the interpretation of the coefficients: while logistic regression reflects changes in log-odds, the probit model interprets coefficients through the lens of standard normal distribution. Despite numerical differences, both models converge on the same conclusion—regional market sales are the dominant indicators of a game's global ranking status.

Table 4. Binary Probit Regression Results for Predicting Above-Average Global Sales Rank

Variable	Coef	Std Err	z	P | z |
Const	-1731.38	8,081,102	0.00	1.000
Year	0.0038	0.0173	0.22	0.826
NA_Sale	86.02	15.79	5.45	0.000
EU_Sale	83.12	15.88	5.24	0.000
JP_Sale	84.09	15.63	5.38	0.000
Other_Sale	91.73	17.88	5.13	0.000
Global_Sale	10051.96	47,535,890	0.00	1.000

Summary of Results

The results from four analytical approaches—correlation analysis, multiple linear regression, binary logistic regression, and binary probit regression—collectively demonstrate consistent statistical relationships across the dataset of global video game sales.

The correlation analysis reveals that NA_Sales and EU_Sales have the strongest linear relationships with Global_Sales, while JP_Sales and Other_Sales show weaker yet still positive correlations. This suggests varying degrees of influence among regional markets on overall sales.

In the multiple linear regression model, regional sales variables exhibit statistically significant negative coefficients, indicating that higher regional sales are associated with better (lower numerical) sales rankings. In contrast, the Global_Sales variable has a positive coefficient, and the Year of release also has a small positive effect on rank. All variables, however, are statistically significant in this model.

The binary logistic regression model confirms that regional sales—NA_Sales, EU_Sales, JP_Sales, and Other_Sales—positively and significantly increase the odds of a video game achieving above-average rank. However, Year and Global_Sales are not statistically significant predictors in the classification context.

Similarly, the binary probit regression model produces nearly identical significance patterns and coefficient directions as the logit model, further supporting the conclusion that regional sales are the most influential predictors of sales rank classification. Once again, Year and Global_Sales are not significant contributors.

Overall, each method provides a unique yet consistent perspective on the data, laying a solid foundation for the interpretive discussion and concluding insights that follow.

DISCUSSION

The analysis presented across correlation, regression, and classification models reveals consistent and interpretable trends in the structure of global video game sales. One of the most prominent patterns is the disproportionate influence of regional sales—particularly from North America and Europe—on global success metrics. This finding not only reinforces the importance of those markets from a revenue perspective, but also suggests that performance in these regions serves as a reliable proxy for overall consumer demand and market acceptance. Japan and other regions, while still influential, exhibit more independent or localized effects, as seen in their weaker correlation coefficients and slightly varied impacts in classification models.

Another noteworthy insight is the limited contribution of the Year variable across all models. While newer games might be expected to benefit from evolving technologies and expanded market access, the data suggests that longevity or release timing alone does not reliably predict higher rank. This could be due to the enduring success of legacy titles or varying marketing strategies independent of release year. Furthermore, the counterintuitive positive coefficient for Global_Sales in the linear model, contrasted with its statistical insignificance in binary models, highlights the complexity of interpreting aggregated variables alongside their components.

Together, these findings emphasize the need for nuanced modeling strategies when analyzing consumer markets as diverse and dynamic as the video game industry. Rather than relying on any single metric or model, a combination of linear, categorical, and probabilistic techniques offers deeper insights into the variables that matter most, and in what contexts they hold predictive value.

CONCLUSION AND FINAL INSIGHTS

This study presents a comprehensive empirical evaluation of the variables that influence global video game sales rankings. Drawing from a robust dataset (13) and employing diverse analytical tools, including correlation analysis (14), multiple regression (15), and probabilistic modeling (9, 16), the research identifies consistent patterns that underscore the critical role of regional sales, platform preferences, and genre alignments.

First, strong positive associations between North American and European sales and global performance demonstrate that success in these markets substantially boosts a game's international ranking (1, 17). These findings confirm that high-performing regions serve as effective indicators of broader market success.

Second, the agreement across linear regression, logit, and probit models reinforces the robustness of the analysis (9, 13, 18). These results show that a combination of statistical modeling approaches can yield more reliable insights into market behavior than reliance on a single method.

Third, the year of release was shown not to be a reliable standalone predictor of performance. While technological improvements and temporal market shifts occur, they do not significantly influence rankings once other variables are accounted for (3). This indicates that recency alone does not determine a game's success.

Future research can expand on these findings by incorporating real-time player sentiment from platforms like Steam and Reddit (19), or by analyzing social and cultural context using qualitative methods (20). Techniques such as clustering and deep learning may uncover nonlinear relationships not addressed by the current models (18, 20).

In conclusion, understanding what drives video game sales rankings involves recognizing where, how, and why games succeed. These insights offer valuable guidance for developers, marketers, and other industry stakeholders aiming to implement data-informed global strategies.

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to the contributors of the Kaggle datasets used in this study. The richness and diversity of this data laid a solid foundation for our analysis and significantly deepened our understanding of the factors that affect different markets for different games. I would also like to express my sincere gratitude to the anonymous reviewers and editors for their valuable contributions to my work.

REFERENCE

1. Wijman T. (2020). Newzoo Global Games Market Report. Available from: https://newzoo.com/resources/trend-reports/newzoo-global-games-market-report-2020-light-version (accessed on 2025-04-29)
2. Beattie A. (n.d.). How the video game industry is changing. Available from: https://www.investopedia.com (accessed on 2025-05-29)
3. Gregor UT. (2023). Video Game Sales Dataset. Available from: https://www.kaggle.com/datasets/gregorut/videogamesales (accessed on 2025-05-29). https://doi.org/10.69849/revistaft/ra10202505231805
4. Marchand A & Hennig-Thurau T. (2013). Value creation in the video game industry. *Journal of Interactive Marketing*. 2013; 27 (3): 141–157. https://doi.org/10.1016/j.intmar.2013.05.001
5. Hamari J & Keronen L. Why do people buy virtual goods: A meta-analysis. *Computers in Human Behavior*. 2017; 71: 59–69. https://doi.org/10.1016/j.chb.2017.01.042
6. Hunicke R, LeBlanc M & Zubek R. MDA: A formal approach to game design and game research. In Proceedings of the AAAI Workshop on Challenges in Game AI. 2004.
7. Kline S, Dyer-Witheford N & De Peuter G. Digital play: The interaction of technology, culture, and marketing. *McGill-Queen's University Press*. 2003. https://doi.org/10.1515/9780773571068
8. Kim Y & Ross SD. An exploratory study of Korean gamers' perceptions of MMORPGs. *International Journal of Computer Games Technology*. 2006.
9. Harrell FE. Regression modeling strategies. Springer. 2015. https://doi.org/10.1007/978-3-319-19425-7
10. Sharpe D. Chi-square test is statistically significant: Now what? Practical Assessment, Research, and Evaluation. 2015; 20 (1): 8.
11. Newzoo. Top Countries & Markets by Game Revenues. Available from: https://newzoo.com/resources/rankings/top-10-countries-by-game-revenues (accessed on 2025-05-29). 2023.
12. Wang H & Sun CT. Game reward systems: Gaming experiences and social meanings. *DiGRA Conference*.

2011. https://doi.org/10.26503/dl.v2011i1.594
13. Entertainment Software Association. 2020 Essential Facts About the Video Game Industry. 2020.
14. Suh A & Wagner C. How gamification of an enterprise collaboration system increases knowledge contribution. *Journal of Knowledge Management*. 2017. https://doi.org/10.1108/JKM-10-2016-0429
15. Vorderer P, Klimmt C & Ritterfeld U. Enjoyment: At the heart of media entertainment. *Communication Theory*. 2004; 14 (4): 388–408. https://doi.org/10.1111/j.1468-2885.2004.tb00321.x
16. James G, Witten D, Hastie T & Tibshirani R. An Introduction to Statistical Learning. *Springer*. 2013. https://doi.org/10.1007/978-1-4614-7138-7
17. Anable A. Playing with feelings: Video games and affect. *University of Minnesota Press*. 2018. https://doi.org/10.5749/j.ctt20mvgwg
18. Statista. Leading video game genres worldwide. Available from: https://www.statista.com/statistics/189592/best-selling-video-game-genres-since-2006/ (accessed on 2025-05-29). 2022.
19. Cohen J, *et al*. Applied multiple regression/correlation analysis for the behavioral sciences. *Routledge*. 2003.
20. Steam DB. Steam Game Charts by Sales. Available from: https://steamdb.info/charts/ (accessed on 2025-05-29). 2023.