Uncovering Sales Patterns in the Video Game Industry

By Kacie Chong, Sumaiya Alvi, Sami Gibson, Hannah Karni, and Praveen Arunshankar

PSTAT 100

Spring 2025

Abstract

Our study analyzed a sample of 500 video games from a Kaggle dataset to investigate factors influencing global and regional video game sales. Three research questions guided the analysis: (1) whether Critic_Score or User_Score better predicts global sales, (2) if specific genres outperform others in North America and Japan, and (3) whether global sales differ significantly across gaming platforms. Results showed that Critic_Score has a stronger and more consistently positive relationship with global sales compared to User_Score. Genre analysis revealed no significant sales differences for Action/Shooter games in North America as suspected, but Role-Playing Games (RPG) significantly outperformed other genres in Japan. Lastly, we confirmed significant differences in global sales across platforms, with platforms like PlayStation 2 and Nintendo 3DS showing notably higher sales. These findings suggest that developers and marketers should prioritize favorable critic reviews, tailor specific genre offerings to regional markets, and strategically select high-selling platforms for game releases. Our limitations included outdated data, broad genre categories, and unaccounted variables, so we recommend that future research incorporate more recent data and multivariate modeling.

Introduction

The Kaggle dataset used in this project, titled "Video Game Sales with Ratings," contains detailed information on over 16,000 video games released globally between 1980 and 2016. It includes both categorical and quantitative variables such as global and regional sales, critic and user scores, platform, genre, and year of release. Our analysis was driven by the following three research questions:

- *Question One:* Does user score or critic score have a stronger relationship with global sales?
- *Question Two:* Do Action and Shooter games have significantly higher average sales in North America, while Role-Playing games have significantly higher average sales in Japan, compared to other genres?
- *Question Three:* Are global sales significantly different across platforms?

We asked these questions to better understand the factors that drive video game sales across different markets. First, we hypothesized that critic scores have a stronger positive correlation with global sales than user scores since professional reviews often shape consumer behavior more directly. Knowing this, companies can mainly prioritize securing strong critic reviews to boost game sales. Second, we explored regional genre preferences, expecting that Action and Shooter games would perform better in North America, while Role-Playing games

would be more popular in Japan, reflecting long-standing cultural trends. With this knowledge, game publishers can tailor their marketing strategies to better meet the tastes of each regional audience. Finally, we hypothesized that global sales would differ significantly across gaming platforms, as platform-specific sales can vary due to factors like brand loyalty and exclusive, popular games specific to a console.

Data Processing

Following the professor's instructions to meet project requirements and maintain a manageable dataset size, we randomly sampled 500 rows from the dataset that had no missing values in any column. Rather than indiscriminately removing incomplete rows, we first assessed the extent and patterns of missingness. Since missing values were minimal, rows containing any missing data were removed. Some missing values were represented as blank strings, which we converted to NA for consistency, while others were explicitly labeled with the string "N/A". This sample kept the diversity of the full dataset while remaining suitable for statistical testing and visualization.

Next, we computed summary statistics for key numerical variables, including release year, regional and global sales, and review scores. This initial exploratory analysis provided insight into data distribution, central tendency, and scale, informing model selection and the need for transformations.

Although some assumptions required for classical linear regression were not fully met, we included a table of summary statistics for key variables to provide essential context for the data. This table helps understand the data distributions and identify potential outliers and skewness.

Table 1: Summary Statistics for Selected Variables

Variable	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Year_of_Release	1996.00	2004.00	2008.00	2007.66600	2011.00	2016.00
NA_Sales	0.00	0.05	0.15	0.32488	0.36	3.66
EU_Sales	0.00	0.02	0.07	0.19300	0.21	3.14
JP_Sales	0.00	0.00	0.00	0.05980	0.02	2.47
Other_Sales	0.00	0.01	0.02	0.08296	0.07	7.53
Global_Sales	0.01	0.11	0.31	0.66074	0.75	11.66
Critic_Score	20.00	62.00	73.00	70.70800	81.00	96.00
Critic_Count	4.00	15.00	25.50	29.84600	42.00	107.00
User_Score	1.90	6.60	7.65	7.26840	8.30	9.50
User_Count	4.00	13.00	31.00	190.51600	105.00	10665.00

Figure 1: Summary Statistics for Key Variables

Transformations

For Question 1, due to the failure of normality assumptions, particularly for the Global_Sales variable, a log transformation was applied to reduce skewness and stabilize variance. To avoid issues with log(0), we used log(Global_Sales + 1). To, again, address skewness in the sales data for Question 2, we applied a log transformation to the regional sales variables. This helped normalize the distribution, making it easier to model the impact of genre groupings on sales within each region. The boxplots compared log-transformed regional sales, in millions, by genre group, which more clearly display differences by spreading out compressed

values and mitigating the influence of extreme outliers. For Question 3, the Global_Sales variable was heavily right-skewed, as proven during the assumption checks. Due to this skewness, plotting raw sales on a linear scale would compress most observations near zero, masking differences among platforms. To improve visualization, we similarly applied a log transformation to the y-axis of the boxplot. This spread out compressed values, reduced the effect of extreme outliers, and made comparisons between the different platforms more visible.

Modeling Process

Question One:

We conducted Pearson correlation analysis and simple linear regression for Question 1. Critic_Score and User_Score were selected as predictor variables based on the hypothesis that professional reviews may influence consumer purchasing behavior more strongly than user reviews. Since all variables were continuous, Pearson correlation coefficients allowed us to directly assess the strength and direction of linear relationships between each score and global sales. This provided an initial understanding of which score might have a stronger association. Next, we implemented separate simple linear regression models for each score to estimate their individual effects on global sales. These models quantified the expected change in global sales for each one-unit increase in the predictor variable. We evaluated model fit through the R-squared value, which indicates the proportion of variance in global sales explained by each score and assessed the statistical significance of predictors using the p-values. Additionally, scatter plots with fitted trend lines were used to visually inspect for potential non-linear patterns and the strength of the correlation. The following hypotheses were tested:

- H_0: The score has no relationship with global sales.
- H_1: The score has a statistically significant relationship with global sales.

Question Two:

We compared average regional sales across genre categories in North America and Japan. Because we were comparing the means of a continuous variable (sales) across two categorical groups, we used Welch's two-sample t-tests, which are appropriate when group sizes and variances are unequal. We also applied simple linear regression models, which allowed us to estimate the direction and magnitude of the relationship between genre groupings and log-transformed regional sales, which served as the dependent variable. We visualized the distributions using boxplots and histograms. We defined region-specific genre groups, which served as the independent variable, based on known cultural preferences. These binary genre variables were used in both the t-tests and regression models:

- For North America, we compared Action/Shooter games against all other genres.
- For Japan, we compared Role-Playing games against all other genres.

The hypotheses tested were:

- North America:

    - H_0: The mean North American sales for Action/Shooter games is equal to that of other genres.
    - H_1: The mean North American sales for Action/Shooter games differs from that of other genres.
- Japan:
    - H_0: The mean Japanese sales for Role-Playing games is equal to that of other genres.
    - H_1: The mean Japanese sales for Role-Playing games differs from that of other genres.

Question Three:

We treated Platform as a categorical predictor and Global_Sales as the continuous response variable; therefore, regression was not appropriate and was not performed. Instead, since we were comparing the means across more than two independent groups, we initially applied a one-way ANOVA test. This test is appropriate for assessing whether any significant differences exist among group means. The hypotheses tested were:

- H_0: The mean Global_Sales is the same across all platforms.
- H_1: At least one platform's mean Global_Sales differs from the others.

To identify which specific platforms differed, we followed up with a Tukey Honestly Significant Difference post-hoc test, which performed pairwise comparisons while controlling for Type I errors. However, after checking for the violations of ANOVA assumptions, a non-parametric alternative was found to be more appropriate. As a result, we turned to the Kruskal-Wallis rank sum test, which does not require normality or equal variances but still allows us to assess whether the median global sales differ significantly across platforms.

Results

Assumption Checks:

Across all three analyses, we conducted thorough assumption checks to ensure the validity of our statistical methods. All models exhibited the same assumption outcomes: deviations from normality and violations of homoscedasticity. However, the independence of observations was assumed satisfied throughout, given the dataset structure where each row represents a unique game. To avoid redundancy as per the professor's instructions, we will present detailed assumption-checking charts and results for only Question 3 below:

To assess the normality of residuals, we examined a histogram, which revealed a strong right skew, deviating from the expected bell-shaped curve. The Q-Q plot showed substantial departures from the reference line, especially in the tails. The Shapiro-Wilk test returned a p-value of 0, which is greater than the significance level of 0.05; therefore, we can strongly reject the assumption of normally distributed residuals.

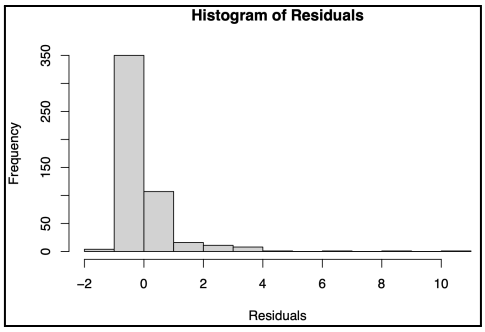

Figure 2: Histogram of Residuals

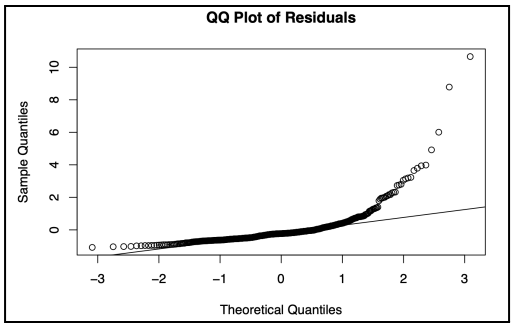

Figure 3: QQ Plot of Residuals

Shapiro-Wilk Test for Normality of Residuals

	Statistic	P_Value
W	0.6033497	0

Figure 4: Shapiro-Wilk Test Results

To evaluate the independence of observations, we plotted residuals against the order of observations. The absence of any clear pattern suggests that the independence assumption was reasonably satisfied. In addition, independence was assumed based on the structure of the dataset, where each row corresponds to a unique video game.

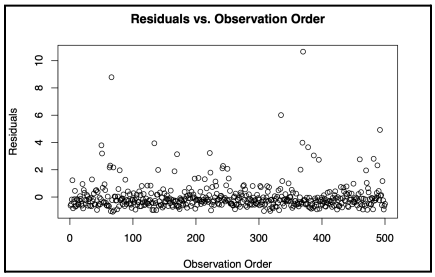

Figure 5: Residuals vs. Observation Order Plot

Levene's Test gave a p-value of 0.095, which is above the significance level of 0.05; we do not have enough evidence to reject the equal variances assumption. The residuals vs. fitted values plot shows increasing spread as fitted values increase, indicating heteroscedasticity.

Levene's Test for Equality of Variances across Platforms

	Df	F value	Pr(>F)
group	16	1.498566	0.0953931
	483	NA	NA

Figure 6: Levene's Test Results

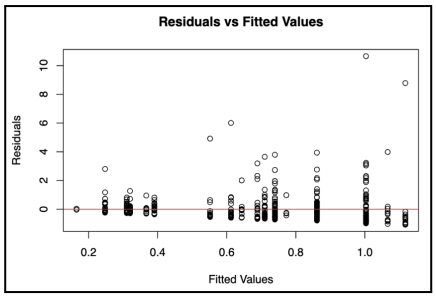

Figure 7: Residuals vs. Fitted Values Plot

Due to the large sample size of 500, linear regression can work with mild heteroscedasticity and the effects of non-normal residuals are minimal. Therefore, it is reasonable to proceed with the analysis while acknowledging these potential limitations. However, to improve model fit and address issues related to skewness, we applied log transformations as mentioned above in the Transformations section.

Question One Results:

We found that the correlation between Critic_Score and Log(Global_Sales + 1) was r = 0.392, indicating a moderate positive relationship. In contrast, User_Score showed a weaker correlation at r = 0.216.

Table 2: Correlation between Scores and Log(Global Sales + 1)

Score_Type	Correlation_with_Log_Global_Sales
Critic Score	0.392
User Score	0.216

Figure 8: Pearson Correlation Between Scores and Log-Transformed Global Sales

To further explore these relationships, we built simple linear regression models. The Critic_Score model had an R^2 of 0.104, which means Critic_Score explains about 10.4% of the variability in global sales. It also has a highly significant p-value, which is less than the significance level of 0.05. This means that we can reject the null hypothesis in favor of the alternative, which states that there is a statistically significant relationship. The regression coefficient was 0.025, indicating that for each one-unit increase in Critic Score, global sales are expected to increase by approximately 0.025 million units, on average. The User_Score model had a lower R^2 of 0.029 and a p-value of 0.00013. While this is still statistically significant, the relationship is weaker and explains only 2.9% of the variation in global sales. The regression coefficient was 0.12, suggesting a smaller and less consistent association.

Table 3: Linear Regression Summary: Global Sales ~ Critic Score

term	estimate	std.error	statistic	p.value
(Intercept)	-1.1150340	0.2382327	-4.680440	3.7e-06
Critic_Score	0.0251142	0.0033055	7.597799	0.0e+00

Figure 9: Linear Regression Summary of Log-Transformed Global Sales on Critic Score

Table 4: Linear Regression Summary: Global Sales ~ User Score

term	estimate	std.error	statistic	p.value
(Intercept)	-0.2817429	0.2491814	-1.130674	0.2587365
User_Score	0.1296686	0.0336400	3.854595	0.0001311

Figure 10: Linear Regression Summary of Log-Transformed Global Sales on User Score

Question Two Results:

Table 1: Two-Sample Welch's t-Test Results by Region

	Region	Group_Comparison	t_statistic	df	p_val	CI_lower	CI_upper	Mean_Other	Mean_Genre
t...1	North America	Action/Shooter vs Other	0.123	427.9	0.9019	-0.085	0.096	0.327	0.321
t...2	Japan	RPG vs Other	-2.047	53.5	0.0456	-0.144	-0.001	0.053	0.126

Figure 11: Two-Sample Welch's t-Test Results Comparing Mean Sales Between Game Genres

For North America, we tested whether the mean sales of Action/Shooter games differ from those of other genres. The results showed a t-statistic of 0.123 with a p-value of 0.9019. The 95% confidence interval for the difference in means ranged from -0.085 to 0.096, and the mean sales were very similar between Action/Shooter games (0.321) and other genres (0.327). Given the high p-value and the confidence interval containing zero, we fail to reject the null hypothesis. This indicates there is no statistically significant difference in sales between Action/Shooter games and other genres in North America. In contrast, when comparing Role-Playing games' mean sales to other genres in Japan, we got a t-statistic of -2.047 with a p-value of 0.0456. The 95% confidence interval ranged from -0.144 to -0.001, and the mean sales for Role-Playing Games (0.126) were higher than those for other genres (0.053). Since the p-value is below the 0.05 significance level and the confidence interval does not include zero, we reject the null hypothesis. This indicates a statistically significant difference in sales, with Role-Playing Games outperforming other genres in Japan.

Table 5: Summary of NA Model (log10(NA_Sales + 1) ~ Genre_Group_NA)

term	estimate	std.error	statistic	p.value
(Intercept)	0.1022796	0.0089273	11.4569021	0.0000000
Genre_Group_NAOther	-0.0011395	0.0112296	-0.1014733	0.9192156
Model Fit (R^2)	0.0000207	-0.0019873	NA	NA

Figure 12: Summary of the North American model from Regression Analysis

Table 6: Summary of JP Model (log10(JP_Sales + 1) ~ Genre_Group_JP)

term	estimate	std.error	statistic	p.value
(Intercept)	0.0176142	0.0027659	6.368341	0.0000000
Genre_Group_JPRPG	0.0268249	0.0090214	2.973479	0.0030872
Model Fit (R^2)	0.0174445	0.0154715	NA	NA

Figure 13: Summary of the Japan model from Regression Analysis

The regression analysis results above support the t-test findings. The North American model showed no significant effect of genre on sales with a p-value of 0.9192, which is greater than the significance level of 0.05, while the Japan model showed a statistically significant effect with a p-value of 0.0031. These results suggest that our predicted genre influences video game sales in Japan but not in North America. To assess the performance of our regression models, we examined the R^2 values. The North American model had an R^2 of 2.07e-5, indicating that genre explained virtually none of the variation in sales. In contrast, the Japan model had an R^2

of 0.0175, suggesting that genre played a somewhat greater role in predicting sales. These performance metrics support our earlier observation that genre appears to have more relevance to sales patterns in Japan than in North America.

Question 3 Results:

Table 1: ANOVA Table: Global Sales by Platform

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Platform	16	37.84889	2.365556	2.063661	0.008872
Residuals	483	553.65853	1.146291	NA	NA

Figure 14: One-way ANOVA Table Results

The one-way ANOVA results revealed a statistically significant effect of the platform on global sales, with a p-value of 0.00887. Since this p-value is below the significance level of 0.05, we reject the null hypothesis and conclude that at least one platform differs significantly from the others in terms of average global sales. To identify which platforms contributed to this difference, a Tukey HSD post-hoc test was conducted, which identified a significant difference between the platforms PlayStation 2 and PC (Personal Computer).

Table 4: Kruskal-Wallis Test Results for Global Sales by Platform

	Statistic	Df	p_value
Kruskal-Wallis chi-squared	86.3297	16	0

Figure 15: Kruskal-Wallis Test Results

Given that the assumptions of ANOVA were violated, the Kruskal-Wallis rank sum test, a non-parametric test, was conducted as an alternative to assess differences in global sales across platforms. The test produced an extremely small p-value, reported as 0 in R due to rounding. This provides very strong evidence against the null hypothesis, confirming that significant differences in global sales exist between platforms.

While traditional performance metrics like RMSE are not applicable to ANOVA or non-parametric tests, the statistical significance and post-hoc Tukey HSD comparisons serve to identify meaningful differences in platform performance. The Kruskal-Wallis test further confirmed these findings with a highly significant result.

Interpretation

Our analysis supports our first hypothesis that Critic_Score more strongly predicts Global Sales compared to User_Score. While both predictors are statistically significant, Critic_Score explains more of the variance in sales and produces a clearer linear trend. This suggests that professional reviews may influence consumer purchasing decisions more reliably than user-generated ratings. These findings answer our research question by providing both statistical and visual evidence that Critic_Score has a stronger relationship with sales performance.

For the second research question, although we initially hypothesized that Action and Shooter games would perform better in North America and Role-Playing Games would dominate in Japan, the analysis revealed that genre has a statistically significant impact on sales in Japan, but not in North America. In North America, Action/Shooter games did not significantly outperform other genres, as evidenced by both the t-test and regression analysis, suggesting that consumer preferences in this region may be driven more by factors such as franchise recognition, marketing, or platform exclusivity rather than genre alone. In contrast, the analysis for Japan demonstrated that Role-Playing Games significantly outperformed other genres, both in mean sales and as a predictor in the regression model.

Lastly, we hypothesized that global sales would differ significantly across gaming platforms, and our analysis supports this claim. The one-way ANOVA results led us to reject the null hypothesis that the mean global sales are equal across all platforms, confirming that at least one platform differs significantly. This was further supported by the Tukey HSD test, which identified a statistically significant difference in sales between PlayStation 2 and PC, highlighting that certain platforms clearly outperform others. Overall, the analysis confirms that global video game sales vary significantly by platform. In particular, the Nintendo 3DS, original PlayStation, and PlayStation 2 exhibited the highest average global sales among the platforms studied, significantly outperforming others such as the Xbox, PC, and Dreamcast.

Visualization and Communication

Question 1 Visualization:

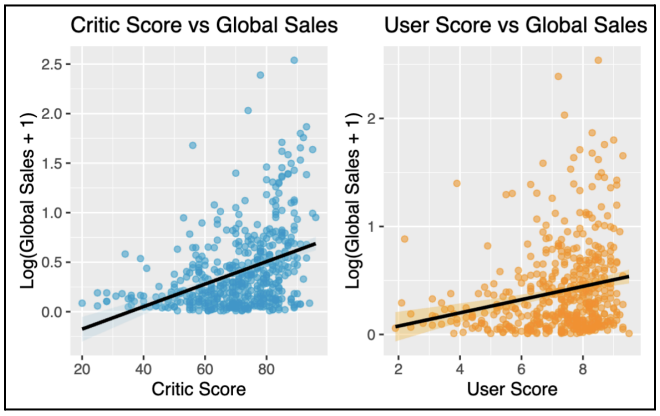

Figure 16: Scatterplots of Critic_Score and User_Score vs. log(Global_Sales+1) with regression lines

The Critic_Score plot shows a clear linear trend and narrow confidence bands, indicating a moderate positive relationship and stronger predictive power for global sales. In contrast, the User_Score plot is more dispersed with a weaker trend and wider confidence bands, suggesting a less consistent relationship. These visualizations support the conclusion that Critic_Score is a more reliable predictor of global video game sales than User_Score.

Question 2 Visualization:

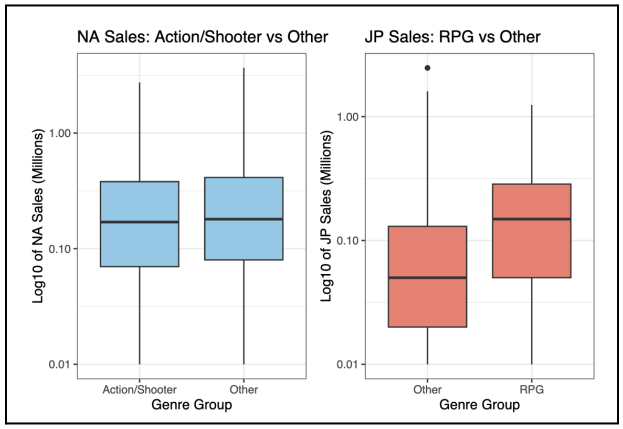

Figure 17: Boxplots of log-transformed regional sales (in millions) by genre group for North America (left) and Japan (right)

The left plot shows North American sales for Action/Shooter games vs. other genres, revealing no outstanding difference in median sales. The right plot shows Japanese sales for Role-Playing Games vs. other genres, indicating higher median sales and a greater upper range for Role-Playing Games. These visuals support the hypothesis that our predicted genre influences sales more significantly in Japan than in North America.

Question 3 Visualization:

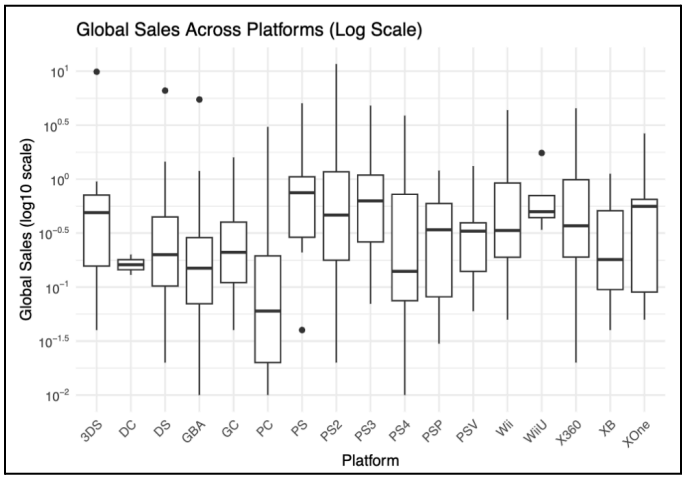

Figure 18: Distribution of Global Video Game Sales by Platform (Log Scale)

Note: The abbreviations for game platforms are used in this figure for brevity. A full list of platform names and their corresponding abbreviations is provided in the appendix.

This boxplot illustrates significant differences in global video game sales across gaming platforms. Medians vary notably: for instance, the original PlayStation, PlayStation 2, and PlayStation 3 show higher median sales compared to platforms like PC, Dreamcast, and

PlayStation 4. The spread of sales also differs: platforms such as PlayStation 4 and PC exhibit wider interquartile ranges, suggesting greater variability. Numerous outliers indicate that some titles achieved exceptional sales. These differences in central tendency and variability provide visual evidence against the assumption of equal medians across all groups. Thus, the plot supports rejecting the null hypothesis, suggesting a statistically significant difference in global sales distributions among platforms.

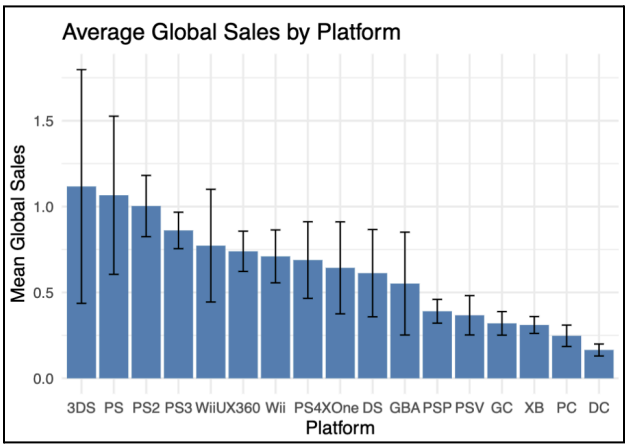

Figure 19: Average Global Video Game Sales by Platform with Standard Error Bars

This bar chart shows significant variation in average global sales across platforms. The 3DS, Sony PlayStation, and PlayStation 2 demonstrate the highest average sales, with the 3DS slightly outperforming the rest. In contrast, platforms such as the Dreamcast, PC, and original Xbox exhibit consistently lower averages. The error bars reveal differences in variability: higher-selling platforms show larger error margins, suggesting the presence of a few top-performing titles. Meanwhile, platforms with lower averages exhibit smaller error bars, indicating more consistent yet modest sales. These visual patterns support the conclusion that platform choice significantly affects a game's commercial success and is helpful for identifying exactly which platforms contribute the most to sales differences.

Conclusion and Recommendations

For Question 1, our findings indicate that both Critic_Score and User_Score are significant predictors of global video game sales, with Critic_Score demonstrating a stronger and more consistent influence. This suggests that developers and marketers should focus on enhancing aspects of the game that resonate with professional reviewers, as positive critic reviews tend to have a clearer and more direct impact on sales performance. Additionally, leveraging favorable critic scores in marketing and promotional strategies can be an effective way to boost consumer interest and drive higher sales.

Question 2 addressed the impact of game genre on regional sales performance, particularly comparing North America and Japan. In Japan, Role-Playing Games significantly outperformed other genres while there was not a notable difference in the genres we tested for North America. Cultural factors, such as a preference for narrative-rich gameplay and the

longstanding popularity of Role-Playing Game franchises, likely contribute to this trend. These insights point to the importance of region-specific marketing strategies. Developers targeting the Japanese market should emphasize Role-Playing Game elements and franchises, whereas a more diversified genre approach may be appropriate in North America.

For Question 3, our analysis revealed statistically significant differences, suggesting that the choice of platform can have a substantial impact on a game's commercial success. Developers and publishers should therefore focus development and promotional resources on platforms with strong sales histories while also investigating underperforming platforms for improvements. For instance, the Nintendo 3DS achieved strong sales likely due to its portability and popular franchises like Pokémon and Mario. In contrast, platforms such as Dreamcast and original Xbox had lower average sales, likely due to shorter lifespans and weaker brand loyalty. These trends suggest that aligning game releases and marketing strategies with specific platforms can help maximize revenue.

Our analyses have several methodological limitations that future research should address. The data used only extends up to 2016, missing more recent platform trends and market changes. Important variables such as marketing budget, franchise popularity, and release timing were not controlled for, though they likely influence sales outcomes. Specifically, the use of broad genre categories in Question 2 limited the analysis by potentially overlooking important distinctions between sub-genres that could differently impact sales. For example, grouping all role-playing games together ignores the differences between action Role-Playing Games and turn-based Role-Playing Games, which can attract different audiences. Future research could benefit from using newer data, breaking genres down into more specific categories, and adding a wider range of predictors. Including variables like release dates, marketing spending, and genre details in multivariate models would help paint a clearer picture of what really drives video game sales. All in all, this project gave us valuable insights into the many factors that drive video game sales such as critic reviews, genre trends, and platform performance, and we strengthened our data science skills through this hands-on experience.

Appendix

Abbreviation	Full Platform Name
3DS	Nintendo 3DS
DC	Sega Dreamcast
DS	Nintendo DS
GBA	Game Boy Advance
GC	Nintendo GameCube
PC	Personal Computer (Windows/Linux)
PS	Sony PlayStation (original)
PS2	Sony PlayStation 2
PS3	Sony PlayStation 3
PS4	Sony PlayStation 4
PSP	PlayStation Portable
PSV	PlayStation Vita
Wii	Nintendo Wii
WiiU	Nintendo Wii U
X360	Xbox 360
XB	Original Xbox
XOne	Xbox One