# Project for Data Science

# Title: Video Game Sales Analysis

| Name |
| --- |
| **Loganesh Raj** |

| | |
| --- | --- |
| **Dataset Link** | https://www.kaggle.com/datasets/ibriiee/video-gamessales-dataset-2022-updated-extra-feat |

# Table of Content

# Introduction

## Problem Statement

*"What are the key factors that influence video game sales performance across time, genre, platforms, regions, and game quality?"*

## Objective

The objective of this analysis is to explore how video game sales have evolved over the past few decades. Using a dataset containing information on game titles, platforms, release years, and regional and global sales figures, this study aims to identify major trends in the industry's growth, peak, and decline phases. Key areas of focus include total global sales, regional distribution patterns, the number of games released per year, and changes in average sales per title over time. The analysis also aims to contextualize these trends within broader market shifts such as the rise of digital distribution .

## Background Information

The global video game industry has evolved into one of the most profitable and influential sectors in entertainment, generating hundreds of billions of dollars in revenue annually. With thousands of games released across various platforms, genres, and regions each year, understanding the factors that drive a game's commercial success has become increasingly important for developers, publishers, and marketers. To understand the dimensions that drives the revenue of video games, this analysis of dataset explores 4 main questions:

1. How do Critics' Scores Correlate with Sales?
2. How have Video Game Sales Trends Evolved Over the Years?
3. How do Sales Figures Vary Across Different Gaming Platforms?
4. How Do Different Video Game Genres Perform Across Various Global Regions in Terms of Sales?

## Scope of the Project

- Geographical scope: North America (NA_Sales), Europe (EU_Sales), Japan (JP_Sales), Other regions(Other_Sales)
- Time period – 2 decades
- Analytical scope – exploratory and correlational

# How do Critic's Scores Correlate with Sales?

In the video game industry, critics' scores are often viewed as a key indicator of a game's, in this case it means high sales. This analysis explores the relationship between critics' scores and game sales to assess whether highly rated scores actually translate into higher commercial performance.

With the use of a cleaned dataset, filtered with the necessary columns, a bunch of statistical analysis has been performed on the dataset to answer the said question with the use of graphs and diagrams.

Figure 1.1 shows a heatmap between the sales of every region recorded in the dataset and the critics' scores.

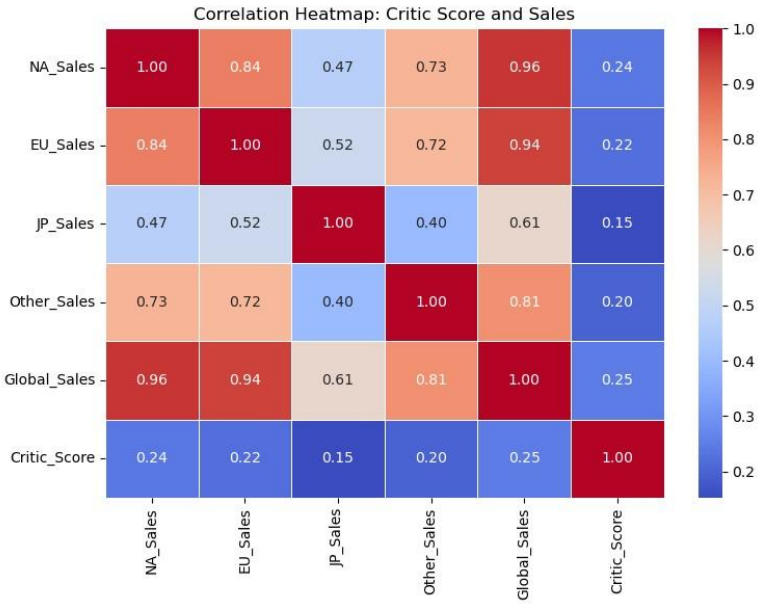


*Figure 1.1*

It is noted that the diagram in Figure 1.1 highlights the fact that all regional sales columns mostly correlate strongly with each other (ranges from 0.4 – 0.96). For instance, correlation between NA_Sales and EU_Sales is 0.84. It can be depicted that success in one region usually mirrors success elsewhere, which is most likely due to brand strength and marketing of the games. On the other hand, every sales column shows a weak but positive (~0.15–0.24) with Critic_Score. Thus, higher critic scores are associated with higher sales, but only slightly. Other variables, such as 'Publisher', 'Platform', 'Genre' etc. may have a bigger impact on sales.

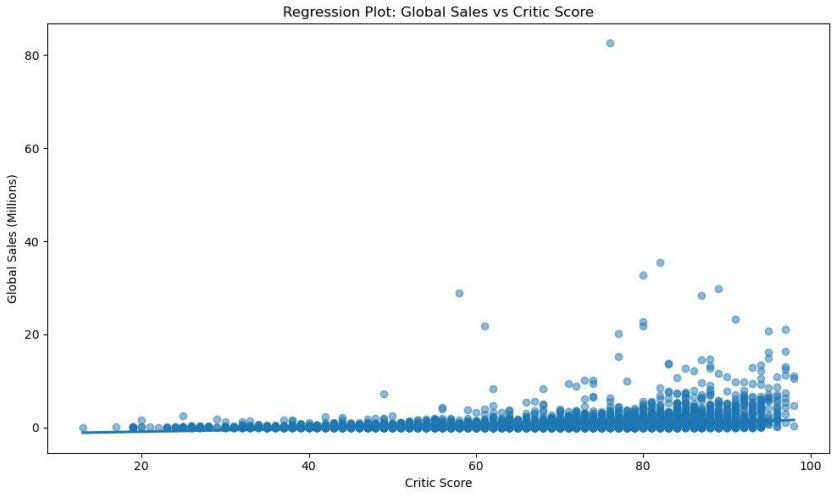


*Figure 1.2*

The regression plot in Figure 1.2 provides additional evidence for the heatmap as it shows slight upward slope, showing that higher critic scores are associated with higher global sales. However, the trend is not strong, with many points scattered far from the line. The correlation coefficient ($r \approx 0.24$) indicates a weak linear relationship. This means critics' scores only explain a small portion of the variation in sales. Adding to the above statement, outliers influence the trend. A few blockbuster games with extremely high sales skew the plot and pull the trend line upward. Without these, the relationship would appear even weaker.

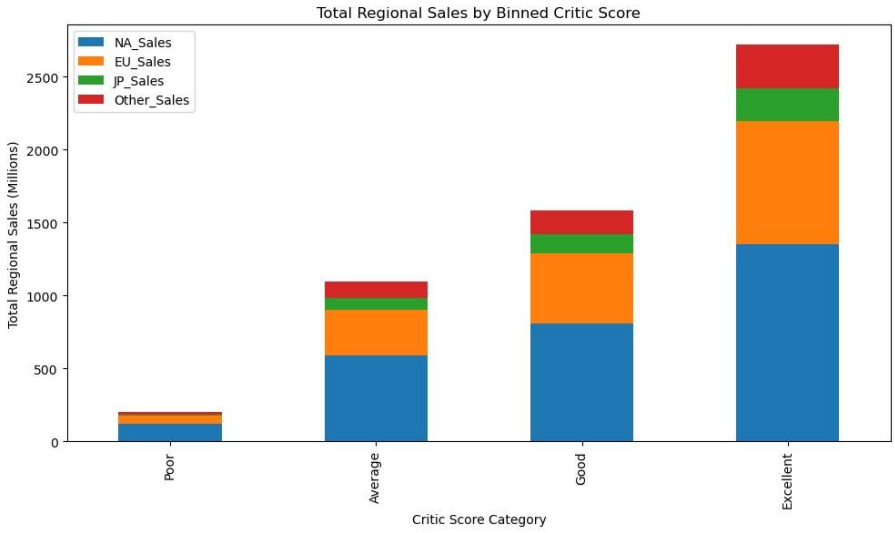


*Figure 1.3*

Figure 1.3 shows a bar chart between Total Regional Sales against the binned critics' scores categories. Here, it conveys that the stacked bars clearly increase in height from Poor to Excellent, indicating that better-reviewed games tend to generate more sales overall. This aligns

with the previous statement that critics' score has a positive, though weak, relationship with sales. Additionally, the Excellent category (Critic Score 85–100) shows the tallest bar, meaning these games collectively earned the most revenue. This suggests that critically acclaimed titles often achieve commercial success. Moreover, in every score category, the largest portions of each stacked bar are from NA_Sales and EU_Sales, indicating these two regions are the main contributors to total global sales.

The boxplot in figure 1.4 below interprets higher critic scores generally lead to higher sales. In all regions, the median sales increase from Poor to Excellent score categories, confirming a positive relationship between critic scores and sales.

North America shows the strongest sales response to critic scores. The highest median and widest spread of sales appear in Good and Excellent games in NA, suggesting this region rewards high-scoring titles more consistently. Oppositely, Japan is least influenced by critic scores. The boxplot for Japan shows minimal changes in sales across score bins, with overlapping distributions. This suggests that critic scores are not a strong predictor of success in Japan, possibly due to cultural or market differences.

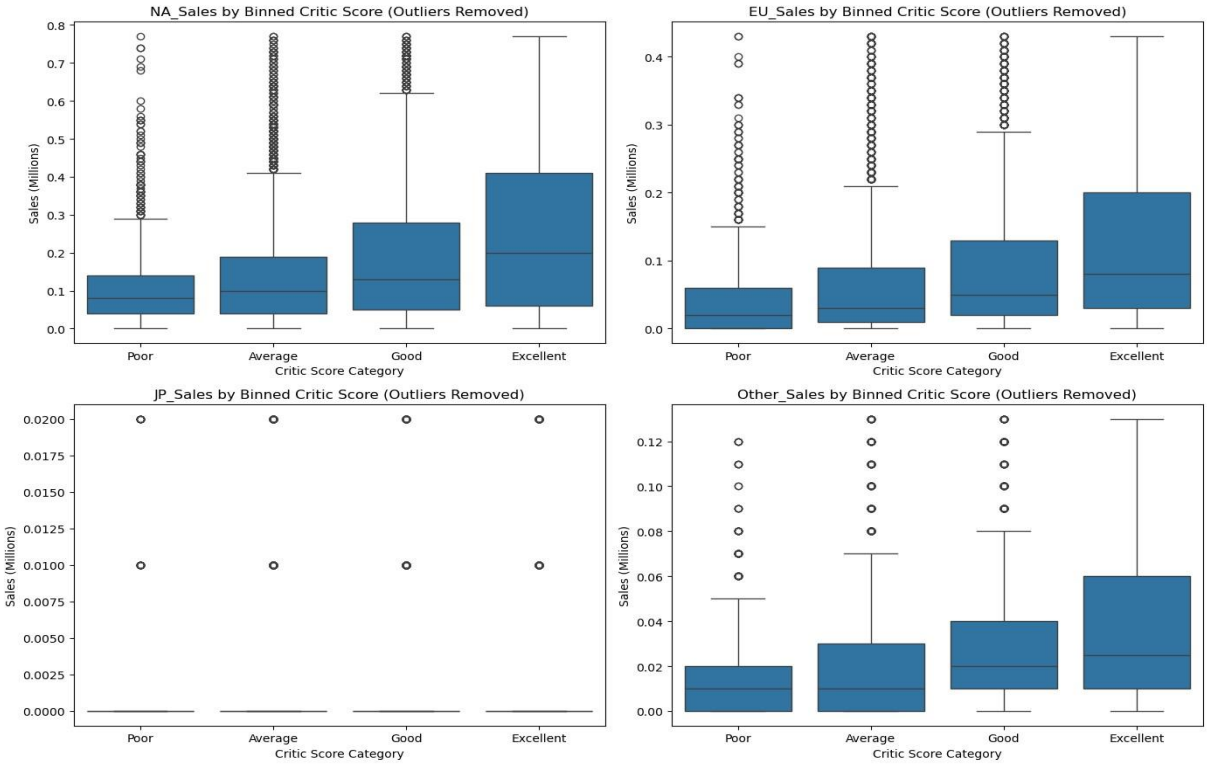


*Figure 1.4*

Limitations:

- After dropping NaNs(rows with missing values), the working sample is cut by around 8100 rows, which is quite a considerable amount, this makes data analysis less credible
- Timing effects are ignored - The dataset bundles 20 years of releases. Average critic scores and sales reporting practices changed dramatically
- Sales are highly skewed A handful of blockbusters (.Mario Kart, .GTA, etc.) sit far to the right of the scatter, inflating the slope and the bin totals.

# How have Video Game Sales Trends Evolved Over the Years?

This analysis explores how global video game sales have changed over time using a cleaned dataset of titles released across multiple platforms. By exploring trends in total sales, regional performance, and the number of games released per year, we aim to recognize the factors behind peaks and troughs in industry performance from the 1980s to the 2020s.

Figure 1 shows a line chart that displays the trends in video game sales throughout the years, with different lines representing sales in various regions. The x-axis represents the year of release, while the y-axis represents the number of sales in millions.

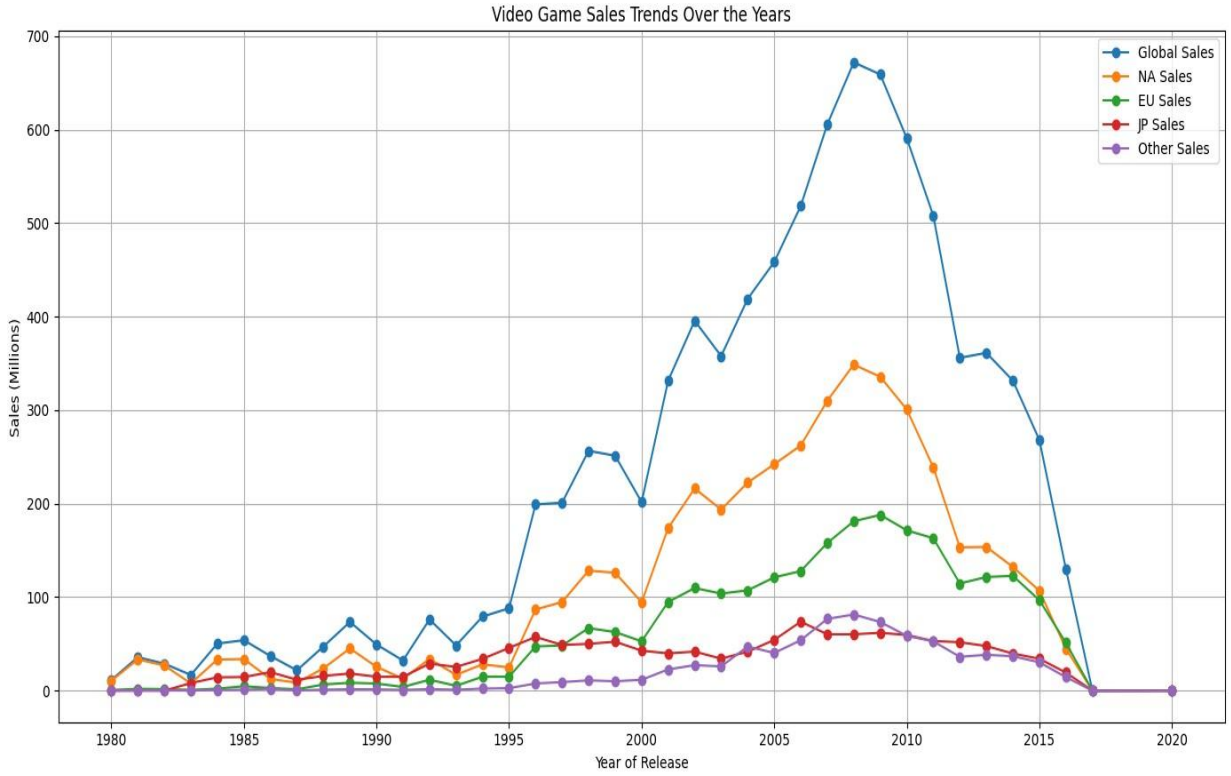


Figure 1: Video Game Sales Trends over the Years

Based on the annual sum shown in Figure 1, global video game sales began increasing significantly from the late 1990s, with the chart showing a striking peak between 2006 and 2009, with international sales reaching nearly 670 million units in 2008. This likely happened because of bigname consoles like Nintendo Wii, PlayStation 3, and Xbox 360, which seems to follow a pattern.

North America and Europe contributed most of the sales across the years, as opposed to Japan, where their sales stayed more consistent over time. Dominated by these big 3, other regions also show gradual growth over time, but remain the smallest market segment.

Around 2010, the market declined drastically in sales, likely due to market saturation or the rise of digital distribution, such as games that can be purchased through PlayStation Store or mobile games.

The sharp drop to near-zero values in recent years likely reflects either incomplete data or the transition away from physical sales tracking, as digital sales became dominant, especially during the pandemic, when physical stores could not be visited due to widespread shutdowns [1].

The chart below displays a bar chart of the number of video games released per year, with each bar representing the count of games for a specific year.

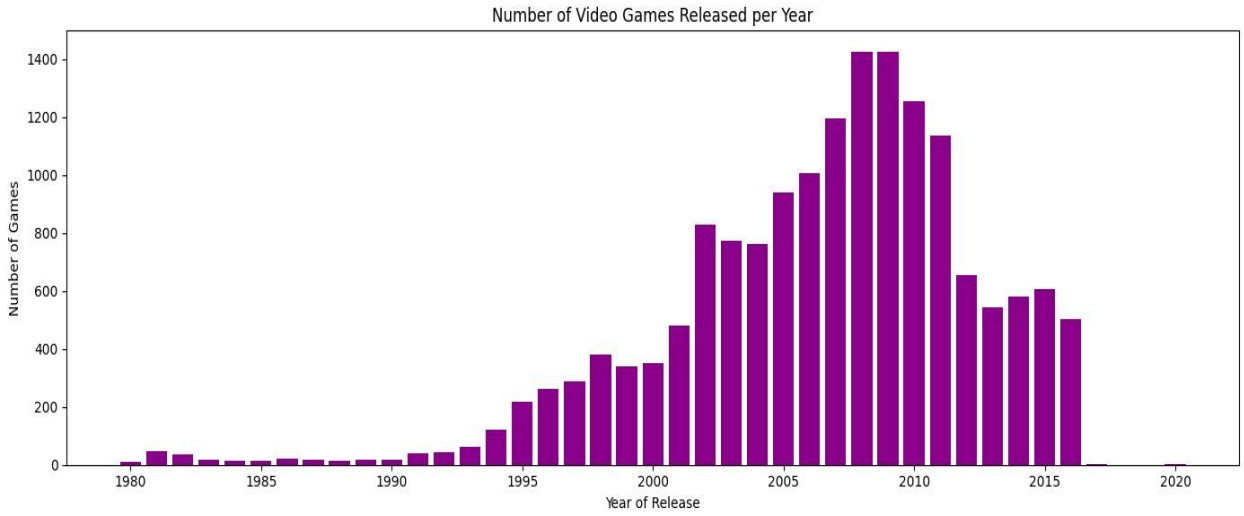


Figure 2: Number of Video Games Released per Year

The highest number of ga releases occurred between 2007 and 2010, with peak years showing approximately 1,400+ releases annually.

After 2010, there is a noticeable decline in the number of releases, dropping to around 500-600 games per year by 2015-2017. This could indicate:

- Market consolidation
- Higher development costs leading to fewer but higher-quality releases

- Shift toward digital-only releases likely missing from the available data • Focus on live-service games and post-launch content

Besides that, figure 2 helps explain what was happening in each era:

- 1980s-1990s: Minimal releases during the early console era
- Mid-1990s: Gradual increase with CD-ROM technology and 3D graphics
- 2000s: Rapid acceleration with multiple competing platforms

The data suggests the traditional retail game market reached peak saturation around 2008-2009, followed by a shift toward digital distribution, mobile gaming, and alternative release models that may not be fully captured in this physical sales dataset.

Figure 3 represents the average global sales per game over the years, with the x-axis represents the year of release, ranging from 1980 to 2020, and the y- axis indicates the average sales (millions).

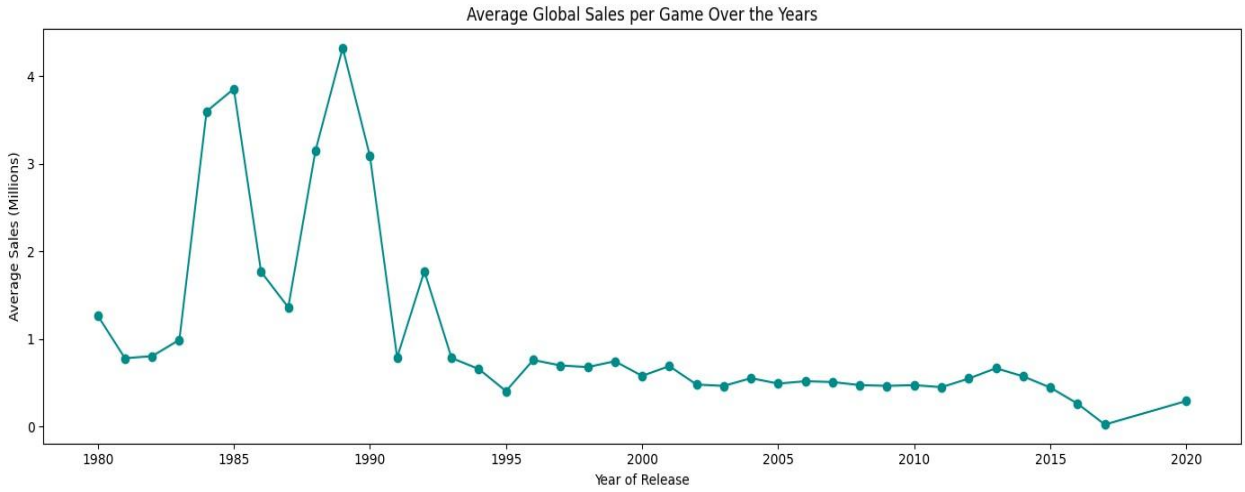


Figure 3: Average Global Sales per Game Over the Years

The chart shows remarkable peaks in average sales per game during certain years, particularly in the mid-1980s and early 1990s, indicating that some games released during those periods had significantly higher sales. This era represents:

- Fewer games released but with higher individual success rates
- Classic console generation (NES era) with less market fragmentation
- Quality-focused development due to higher barriers to entry

After the early 2000s, the average sales per game appear to decline slightly and stabilize at relatively low levels, ranging between 0.3-0.7 million per game, indicating:

- Market saturation with thousands of games released annually
- Long-tail distribution where few blockbusters dominate while many games achieve modest sales
- Rise of niche markets and indie games

The spike drop in recent years, specifically around 2017-2018, likely reflects:

- Incomplete data for digital sales
- Shift toward free-to-play and service-based models not captured in unit sales
- Mobile gaming expansion diluting traditional retail metrics

Back then, releasing a game meant you had a real shot at major success if it hit the right notes. Today, with thousands of new titles every year, the industry feels more like a crowded marketplace. Profit does not come from just selling copies – it is about long-term engagement, smart marketing, and finding the right niche.

# How do Sales Figures Vary Across Different Gaming Platforms?

The performance of gaming platforms varies substantially across different global markets. By exploring platform sales using their global and regional performances, this section aims to uncover the quality and quantity of commercial success for each platform.

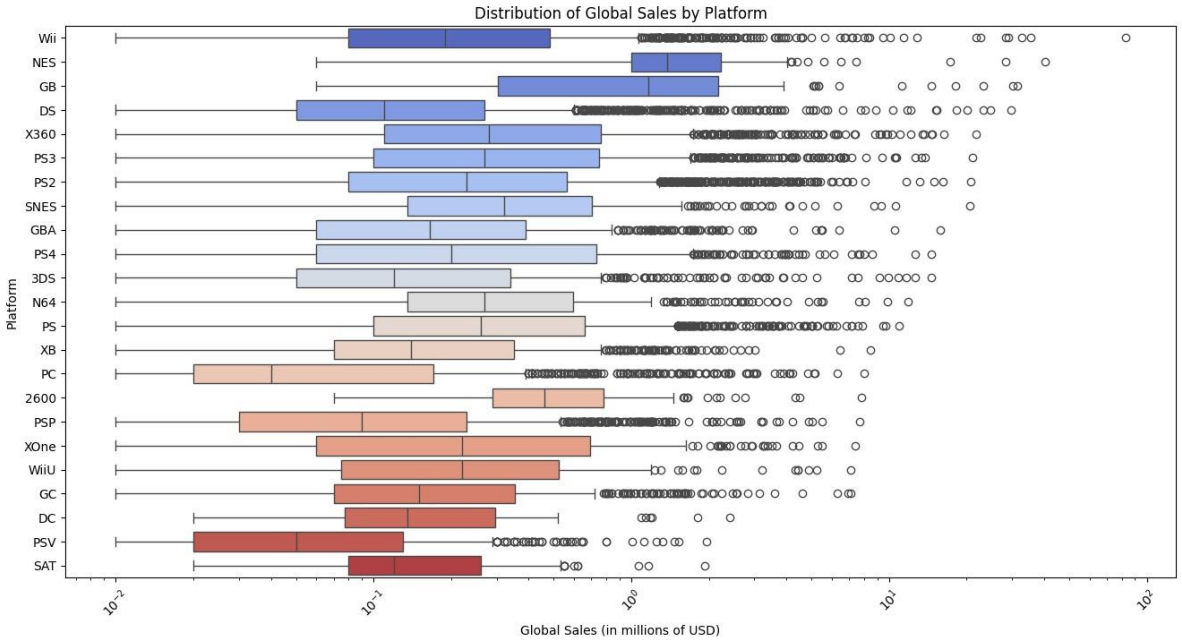


*Figure 3.1: Distribution of Global Sales by Platform*

Based on the box plot shown in Figure 3.1, NES, GB, and Atari 2600 have higher median global sales per game. Wii exhibits significant outliers, indicating large blockbuster titles. Niche platforms like DC, PSV, and SAT show low medians and small spreads, suggesting consistently lower commercial success for games based on those platforms.

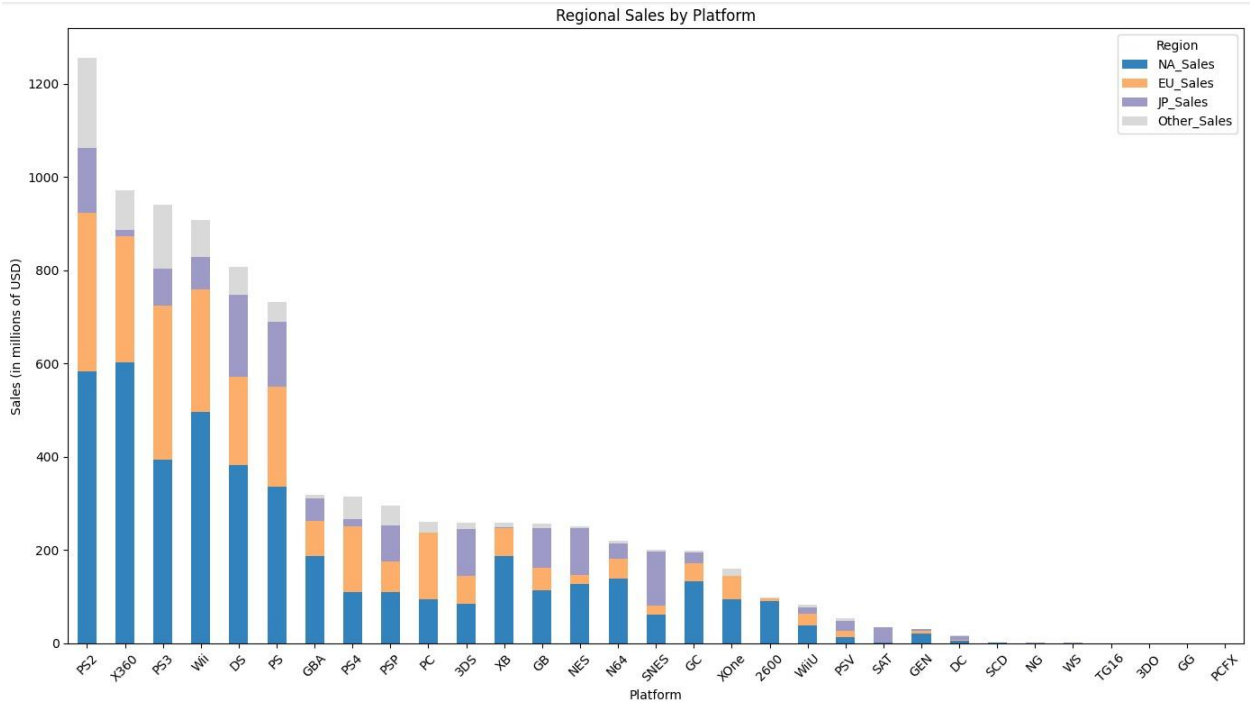


*Figure 3.2: Regional Sales by Platform*

Based on the stacked bar chart shown in Figure 3.2, PS2 leads in total sales across all regions, cementing its position as the most commercially successful platform. X360 show strong performance in North America (NA) and Europe (EU), while Japan-native platforms like PS3, Wii, DS, and PS maintain a more globally balanced performance. NA is the largest market contributor for nearly all platforms, followed by EU.

These findings show us that platform success is multi-dimensional. High global sales can thrive on a few bestsellers, as showcased by Wii and DS, or consistent mid-level performers like PS2, X360, PS3, and PS. Platform country of origin shapes regional dominance, as western platforms like X360 see more success in NA and EU, while Japan-native platforms benefit from strong homemarket support. Certain platforms rely heavily on outliers, suggesting the importance of flagship games and franchises.

These findings can help guide future decision-making for developers and investors alike. Game developers should prioritize platforms with balanced regional appeal like DS and PS for globally oriented games. They should avoid platforms with huge outliers like Wii due to their success concentrating on a few games, mainly Nintendo's own titles that are coupled to the platform (Babb et al., 2013). Investors should be weary of platforms with broad box plot spreads, such as Wii and DS, as they suggest higher market volatility which may affect portfolio risk.

# How Do Different Video Game Genres Perform Across Various Global Regions in Terms of Sales?

The performance of video game genres varies substantially across different global markets. By analyzing sales data broken down by genre and region namely North America (NA), Europe (EU), Japan (JP), and Rest of the World (ROW) this section aims to uncover patterns of regional preference and identify which genres perform best in each market.

| | Genre | NA_Sales | EU_Sales | JP_Sales | ROW_Sales | Global_Sales |
|---|---|---|---|---|---|---|
| 0 | Action | 879.0 | 519.1 | 161.4 | 184.6 | 1744.2 |
| 1 | Sports | 684.4 | 376.8 | 135.5 | 134.5 | 1331.3 |
| 2 | Shooter | 592.2 | 317.3 | 38.8 | 104.1 | 1052.4 |
| 3 | Role-Playing | 330.8 | 188.7 | 355.5 | 59.6 | 934.6 |
| 4 | Platform | 445.5 | 200.4 | 130.8 | 51.1 | 827.8 |
| 5 | Misc | 407.3 | 212.7 | 108.1 | 74.4 | 802.5 |
| 6 | Racing | 359.4 | 236.5 | 56.7 | 76.1 | 728.7 |
| 7 | Fighting | 223.4 | 100.3 | 87.5 | 36.4 | 447.5 |
| 8 | Simulation | 182.2 | 113.5 | 63.8 | 30.8 | 390.3 |
| 9 | Puzzle | 122.9 | 50.0 | 57.3 | 12.4 | 242.6 |
| 10 | Adventure | 105.3 | 63.5 | 52.3 | 16.5 | 237.6 |
| 11 | Strategy | 68.6 | 45.2 | 49.7 | 10.8 | 174.2 |

*Table 1: Total Video Game Sales by Genre and Region (in millions)*

From the aggregated data in Table 1, it is clear that Action games are the leading genre globally, with significantly high sales figures in both North America and Europe. This strong performance can be attributed to the genre's broad appeal, fast paced gameplay, and consistent presence across popular platforms. Sports games follow closely behind, also showing robust performance in Western markets. This trend aligns with the cultural popularity of sports like football and basketball in these regions, which are frequently adapted into video games.

Shooter games also show impressive numbers, particularly in North America and Europe, further emphasizing Western gamers preference for competitive and often multiplayer based titles. In contrast, Japan shows a different trend, where Role Playing Games (RPGs) have a much stronger foothold. RPGs in Japan often feature deep storytelling, character development, and turn based combat mechanics that resonate more with local cultural preferences and long standing gaming traditions.

Genres such as Platform and Miscellaneous display a relatively balanced presence across regions, although their sales are still predominantly influenced by the North American market. Puzzle, Adventure, and Strategy games, while present in all markets, show more modest sales figures globally. These genres tend to have niche audiences and appeal to specific age groups or player types, which may limit their mass market success.

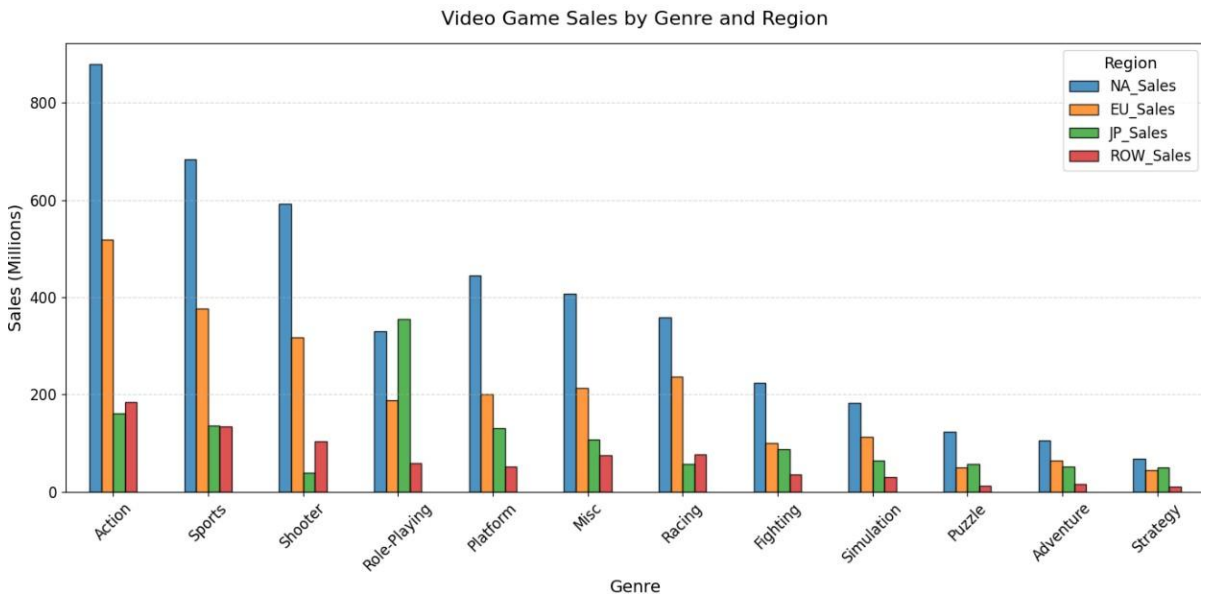


*Figure 4.1: Video Game Sales by Genre and Region*

Figure 4.1 provides a visual representation of total unit sales for each genre across the four major global regions. The chart makes it clear that Action, Sports, and Shooter games dominate sales in North America and Europe, reinforcing the earlier numerical findings. These genres have benefited from years of blockbuster franchises, high advertising budgets, and competitive esports scenes. In comparison, Japan is the only region where Role Playing games significantly outperform all other genres, highlighting a distinct divergence in player interests.

The figure also reveals that Rest of the World regions, while contributing lower total volumes, still maintain a presence across most genres. This suggests that although emerging markets are smaller in scale, they are steadily becoming more involved in global gaming trends.

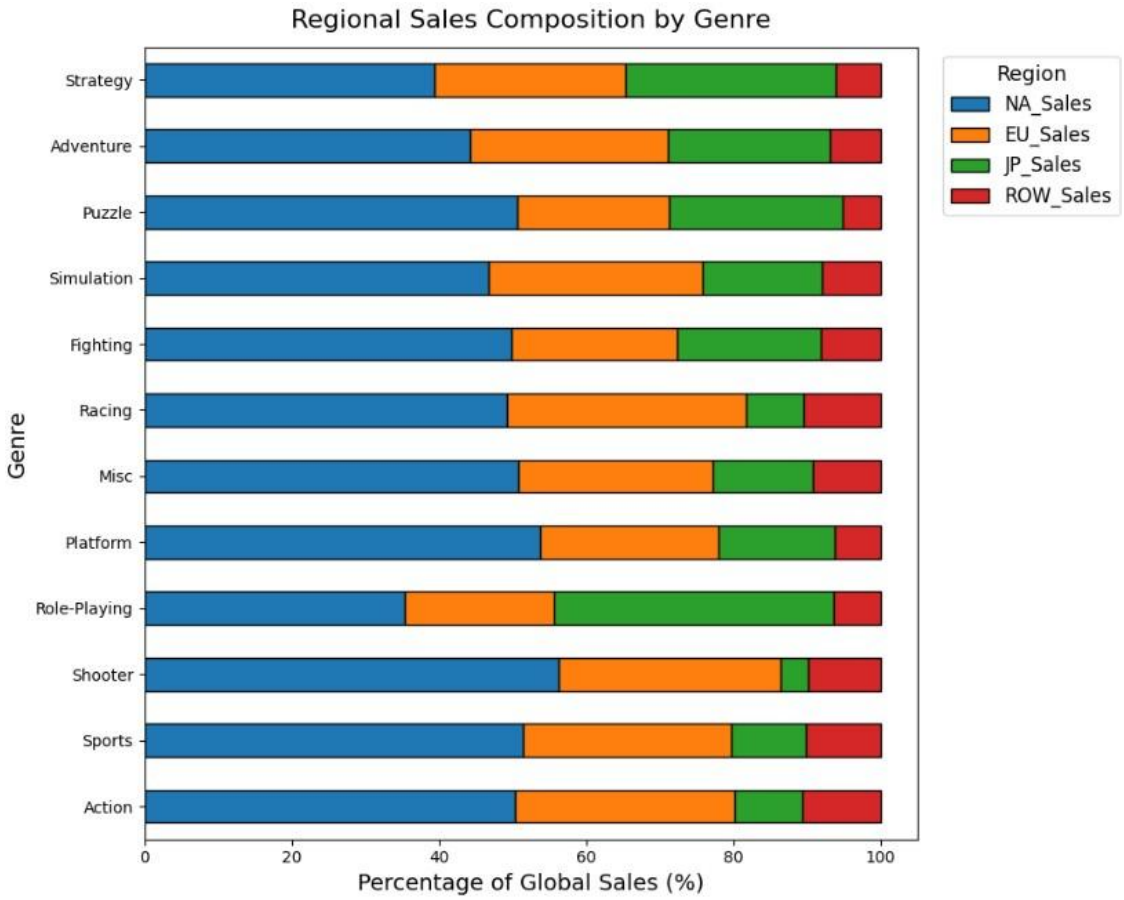


*Figure 4.2: Regional Sales Composition by Genre*

In Figure 4.2, the focus shifts to the percentage composition of each genre's global sales by region. This format helps to understand the relative market importance of each region for a given genre. North America leads in several genres, including Shooter, Action, and Racing, typically contributing between 45% to 55% of total sales in these categories. Europe shows similar patterns, albeit slightly lower, usually around 25% to 35% depending on the genre.

Japan stands out clearly for its dominance in Role Playing games, contributing approximately 38% of global RPG sales, and showing notable participation in Platform and Simulation genres as well. This underlines the importance of localising game content for Japanese audiences, as their preferences differ significantly from Western markets.

ROW regions contribute around 8–15% of global sales depending on the genre. Although their share is smaller, their involvement is consistent across the board, suggesting that growth potential exists in these markets if distribution, localisation, and pricing are appropriately adjusted.

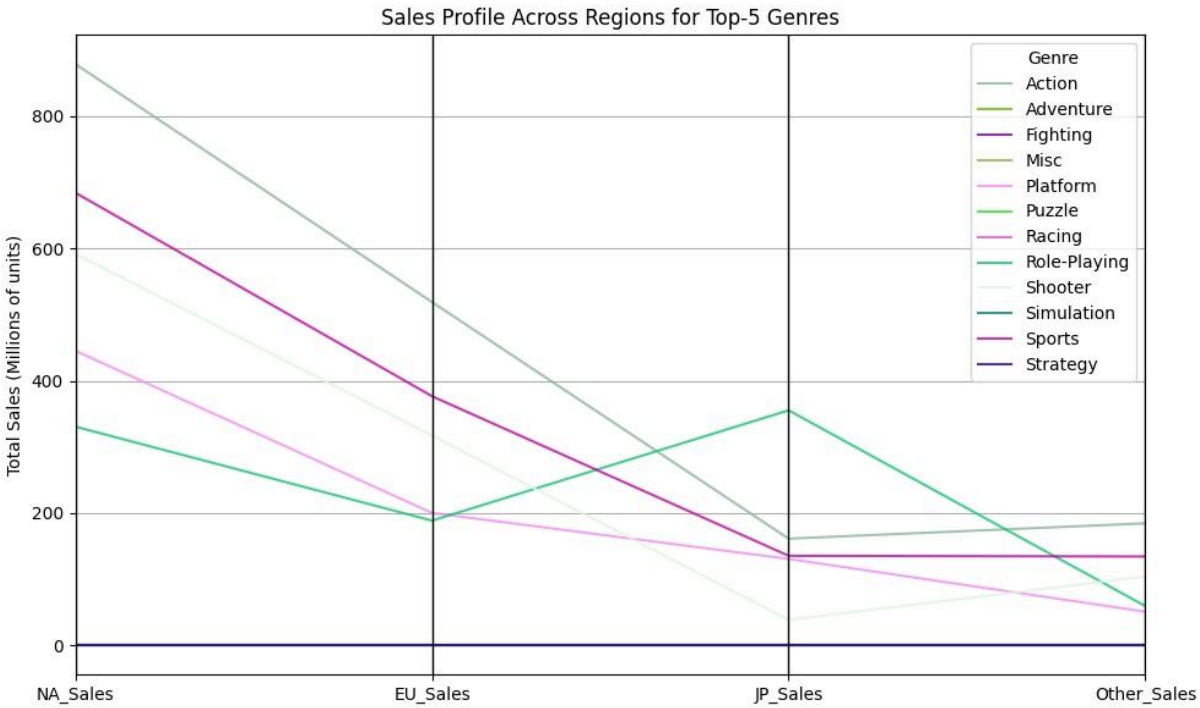


*Figure 4.3: Sales Profile Across Regions for Top-5 Genres*

The parallel coordinates chart in Figure 4.3 focuses on the top five highest selling genres globally: Action, Sports, Shooter, Role Playing, and Platform. Each line represents a genre's sales volume across the four regions. The figure shows that Action games consistently achieve high sales across all regions, affirming their status as a globally appealing genre. Sports games follow a similar trajectory, supported by yearly franchise releases and wide international fandoms.

The Shooter genre shows a sharp decline in Japanese sales, highlighting its limited traction in that region. Conversely, Role-Playing games see their highest sales in Japan, standing out as the only genre in which the Japanese market contributes more than any other. This insight emphasizes the need for culturally tailored game design and marketing strategies when approaching the Japanese gaming audience.

Platform games, meanwhile, appear evenly distributed across regions, suggesting they are wellreceived but less dominant. They may appeal across age groups and platforms but don't command the same market share as the more aggressive genres.

# Conclusion

This project analyzed video game sales data from multiple perspectives, uncovering valuable insights into the evolution of the gaming industry, regional and genre-based preferences, platform dynamics, and the relationship between critic reviews and commercial success.

The sales trend analysis revealed a significant growth phase from the late 1990s to 2008, driven by successful console generations such as the Nintendo Wii, PlayStation 3, and Xbox 360. This era represented the peak of physical game sales, followed by a noticeable decline beginning in 2010. Contributing factors included the shift toward digital distribution, increasing development costs, and changes in consumer behavior. The decline in the number of new releases in recent years reflects industry consolidation and a move toward quality-focused development.

Genre-based performance showed strong regional preferences. Action, Sports, and Shooter games were dominant in North America and Europe, while Role-Playing Games (RPGs) were particularly favored in Japan. This distinction highlights cultural differences in game design preferences and the importance of regional localization. Although Rest of the World regions contributed smaller sales volumes, they demonstrated consistent engagement across most genres, suggesting untapped growth potential.

Platform-based analysis confirmed that consoles, particularly those from Sony, Microsoft, and Nintendo, continue to dominate sales, while PC platforms remain relevant for niche genres and indie games. Limitations in the dataset, particularly regarding mobile and digital-only platforms, suggest the need for future studies with more inclusive data coverage.

A moderate positive correlation was found between critic scores and game sales, indicating that high ratings can enhance commercial performance. However, it is also evident that marketing, franchise loyalty, and platform exclusivity play substantial roles in a game's success.

In summary, this project underscores how global video game sales are shaped by technological evolution, regional preferences, and changing business models. Stakeholders in the industry must consider these factors when designing, marketing, and distributing games to meet diverse player expectations across markets.

## References

[1] Gameopedia. (2023, September 21). *The decline of physical games and the rise of digital distribution*. https://gameopedia.com/blogs/the-decline-of-physical-games-and-the-rise-ofdigital-distribution

Ibriee (2022). Video Game Sales Dataset (Updated with Extra Features). Kaggle. https://www.kaggle.com/datasets/ibriiee/video-games-sales-dataset-2022-updated-extra-feat

Huntemann, N., & Payne, M. T. (2010). Joystick Soldiers: The Politics of Play in Military Video Games. Routledge.

Babb, J., Terry, N., & Dana, K. (2013). The impact of platform on global video game sales. International Business & Economics Research Journal (IBER), 12(10), 1273. https://doi.org/10.19030/iber.v12i10.8136