

# Data Analysis Report on Video Games Dataset

Link: [https://colab.research.google.com/drive/1FoprLQyqbJ3vK10\\_WU4C9cyFN8nHSsfK?usp=sharing](https://colab.research.google.com/drive/1FoprLQyqbJ3vK10_WU4C9cyFN8nHSsfK?usp=sharing)

Name: Akshayaram S

## Introduction:

This report provides an overview of the exploratory data analysis performed on the game sales dataset. The dataset contains information on 1907 different video games, including the game title, platform, release year, genre, publisher, and sales figures in various regions around the world.

The objective of this analysis was to gain a deeper understanding of the video game industry and the factors that contribute to a game's success. Through visualizations and summary statistics, we aimed to uncover patterns and relationships between different variables and identify any trends or key players in the market.

In this report, we will present the results of our analysis and discuss the insights we have gained from the data. By understanding the characteristics of successful games and the trends shaping the video game industry, we hope to provide valuable information for game developers, publishers, and stakeholders in the industry.

## Data Import and Preprocessing:

The data was imported using the Pandas library and read from a .csv file stored in the local storage. The following libraries were also imported for data visualization and manipulation: Numpy and Plotly. The first step in the analysis was to check for missing values in the dataset and handle them appropriately. The missing values in the numerical columns were filled with the median of the column, while the missing values in the categorical columns were filled with the column's mode.

## Data Statistics & EDA

After preprocessing the data, the next step was to perform a descriptive analysis of the dataset, which involved calculating the shape of the dataset, the information contained in each column, and the summary statistics of the numerical features. The shape of the dataset was found to be 1907 rows and 12 columns. The information contained in each column was found to be mostly non-null. The summary statistics of the numerical features showed the distribution of the sales in different regions (North America, Europe, Japan, Rest of World, and Global) and the average review score of the games.

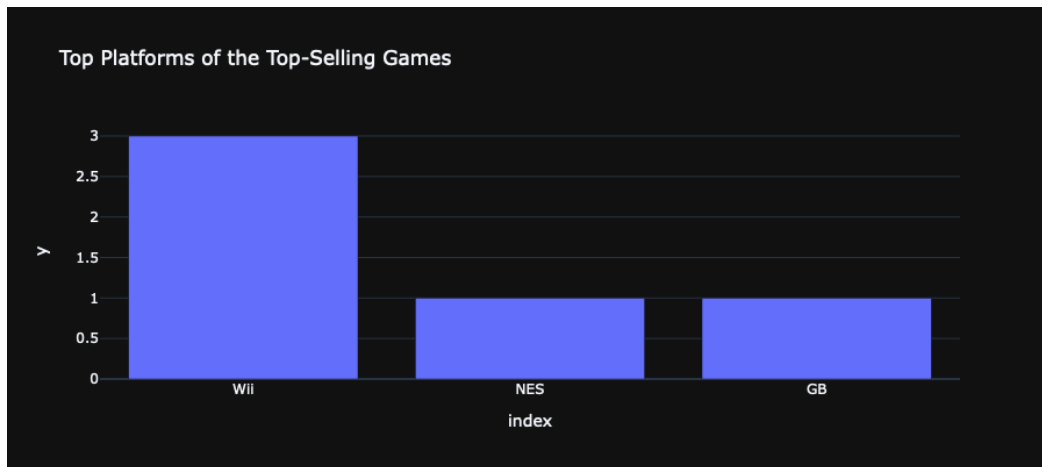


Fig 1:Top Platform of Top-Selling Games

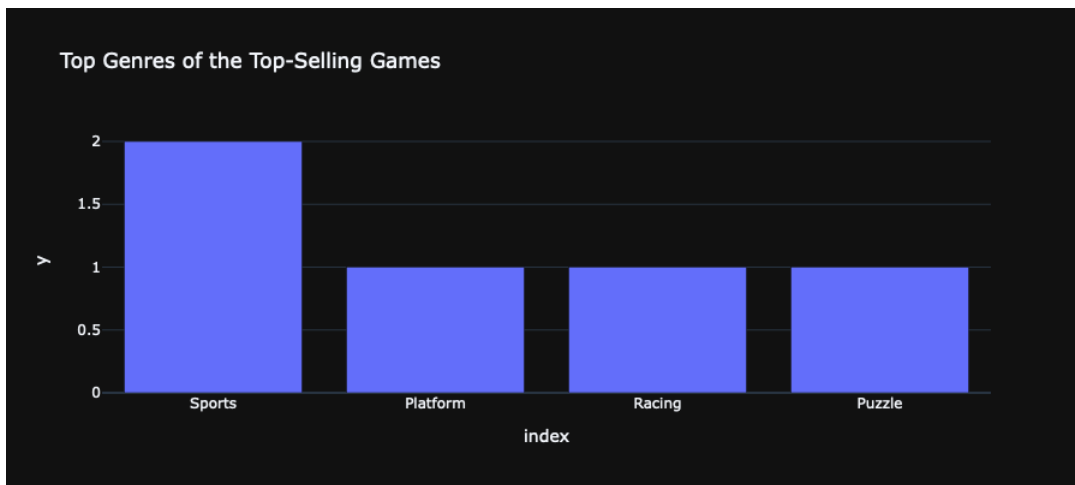


Fig 2: Top Genres of the Top-Selling Games

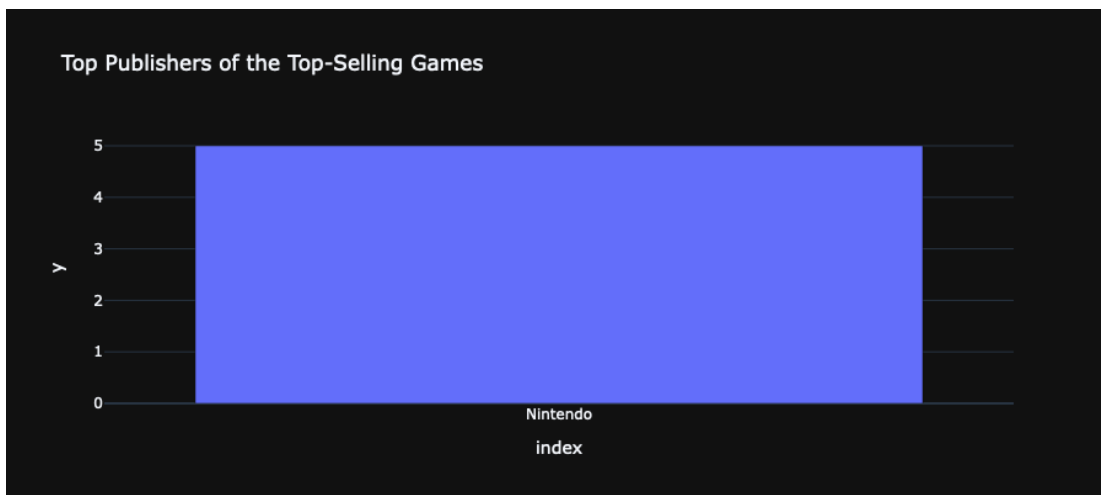


Fig 3: Top Publishers of the Top-Selling Games

Three graphs were generated to gain insights into the top-performing games in terms of sales. The first graph showed the top platforms used by the top-performing games, the second graph showed the top genres among the top-performing games, and the third graph showed the top publishers among the top-performing games.

The top 5 selling games are published by Nintendo, with 5 games in the list. The most common genre among these top games is Sports, followed by Platform, Racing, and Puzzle. The most common platform used among these games is Wii, with 3 games in the list, followed by NES and GB. This suggests that the Nintendo Wii platform, sports genre, and Nintendo publisher have had a strong impact on the success of these top-selling games. Additionally, having a diverse range of genres, such as Sports, Platform, Racing, and Puzzle, has also contributed to the success of these games.

This information provides valuable insights into the factors that contribute to the success of video games and can be used to inform future game development and marketing strategies.

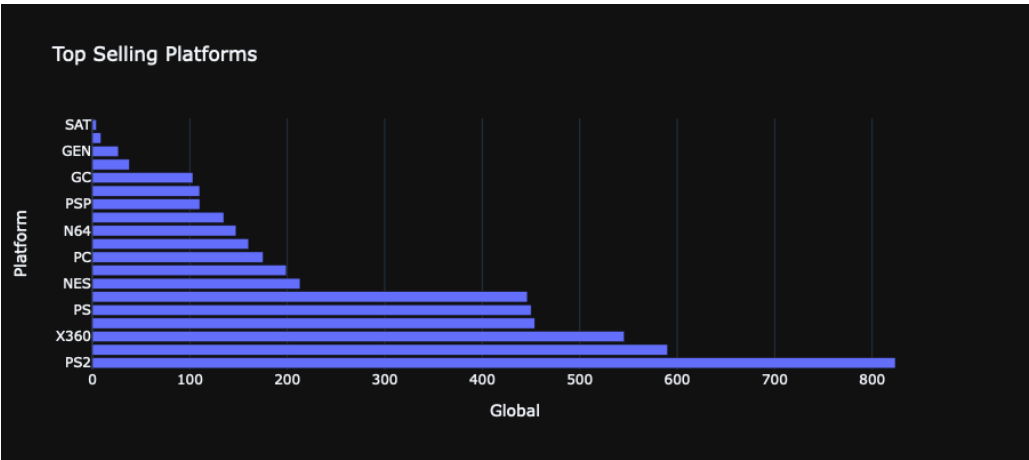


Fig 4: Top Selling Platforms

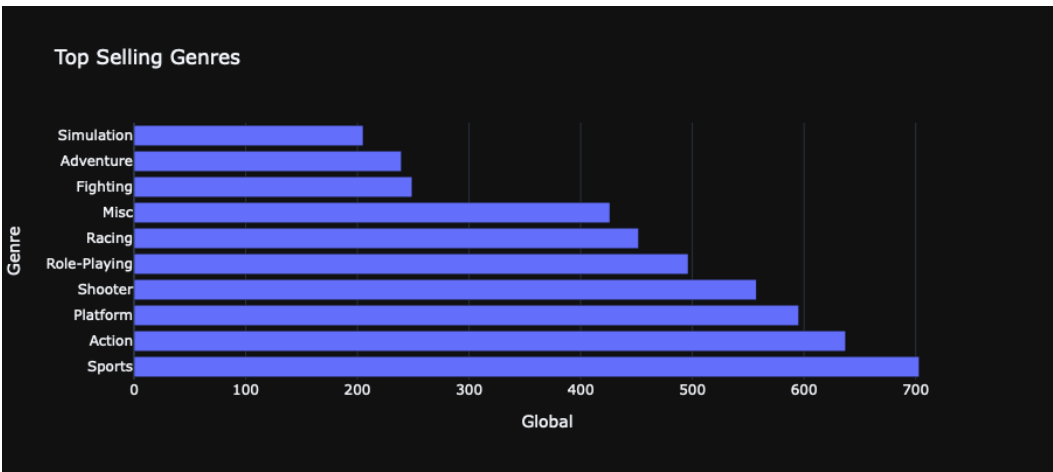
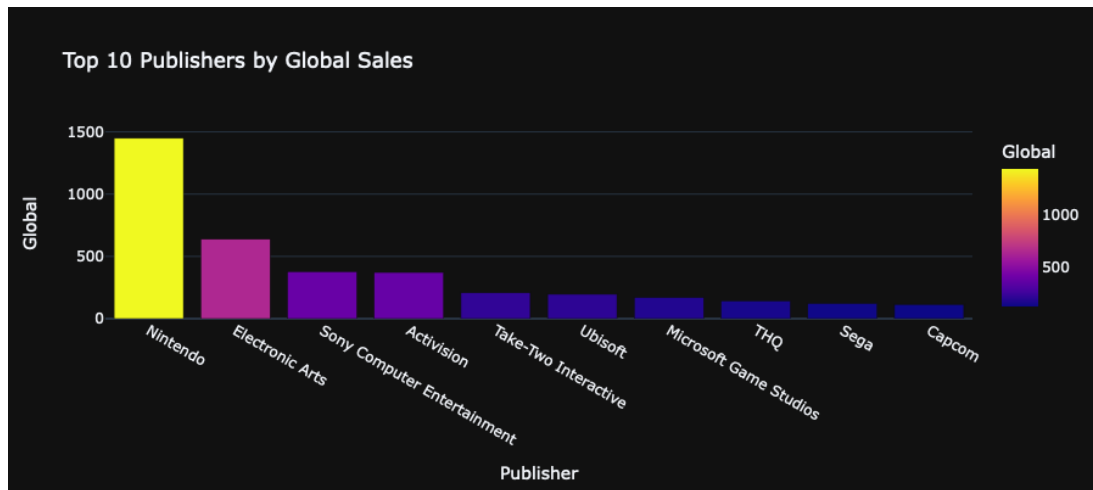


Fig 5: Top Selling Genres

The next step in the analysis was to examine the top platforms and genres by global sales trends. This was done using visualizations such as bar charts.

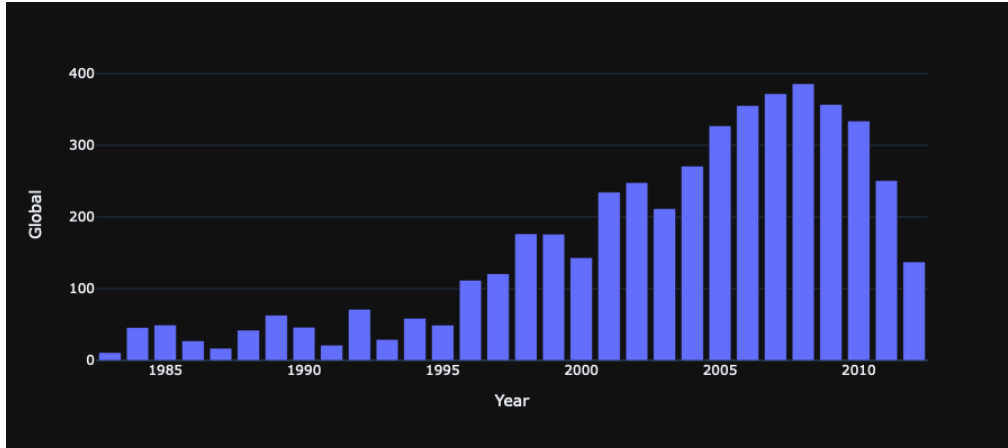
The top 3 platforms with the highest global sales were found to be PS2, Wii, and X360. The genres with the highest global sales were Sports, Action, and Platform. These insights can provide valuable information for game developers when making decisions about which platforms and genres to develop for.

The data suggests that there is a higher demand for Sports and Action games compared to other genres, and developing games for PS2, Wii, and X360 platforms could be a profitable decision. This information can also help game publishers decide which platforms and genres to prioritize in terms of investment and marketing efforts.



**Fig 6: Top 10 Publishers by Global Sales**

Based on the data, it can be seen that Nintendo is the publisher with the highest global sales. They have consistently released successful games and are the top-selling publisher in North America, Europe, and Japan. Electronic Arts come in second place, followed by Sony Computer Entertainment, Activision, and Take-Two Interactive. The data suggests that these publishers have been consistently successful in their game releases and have a large market share. Knowing this information can help other game developers and publishers make informed decisions on which companies to collaborate with or emulate in order to achieve success.

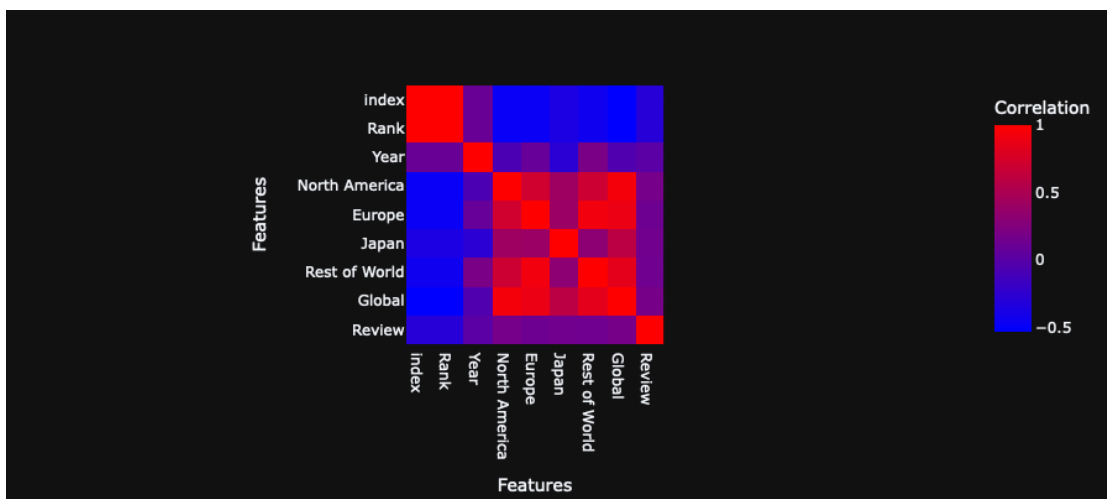


**Fig 7: Global Sales for Each Year**

One of the analyses performed was on the global sales trend of video games over the years. A bar graph was created to visualize the global sales for each year. It was found that there was an overall upward trend in the sales of video games, with some fluctuations. The highest global sales were in the late 90s and early 2000s, with a peak in 2006, before declining somewhat in the later years.

This trend can likely be attributed to the growth and evolution of the video game industry, with new technology and platforms being introduced and adopted, leading to an increase in sales. The decline in later years could be due to the global financial crisis of 2008, which had an impact on consumer spending and the video game industry.

These insights into the global sales trend of video games over the years provide valuable information to game developers and publishers, who can use this information to make informed decisions about their game development and release schedules.



**Fig 8: Correlation Matrix**

The regression analysis showed a strong positive correlation between "Global Sales" and "North America Sales" (0.933073), "Europe Sales" (0.888902), and "Rest of World Sales" (0.837469). This suggests that as sales in one region increase, sales in other regions also tend to increase. The analysis also showed a moderate positive correlation between "North America Sales" and "Europe Sales" (0.720766).

On the other hand, there is a weak positive correlation between "Year" and "Global Sales" (0.201001), which suggests that as the year increases, global sales tend to increase. The analysis also found a weak negative correlation between "Rank" and "Global Sales" (-0.529373), indicating that as the rank increases, global sales tend to decrease, but the correlation is not very strong.

The regression analysis also showed a weak negative correlation between "Review" and "Global Sales" (-0.292892), suggesting that as the review score increases, global sales tend to decrease, but the correlation is not very strong. Finally, the analysis found a weak positive correlation between "Japan Sales" and "Review" (0.148584), indicating that as the review score increases, Japan sales tend to increase, but the correlation is not very strong.

### **Conclusion:**

Based on the analysis of the video game industry sales trends, it can be concluded that the industry has experienced growth over the years, with some fluctuations. Nintendo is the top-selling publisher globally, and the popularity of Sports and Action games has been consistent among players. Additionally, there is a strong positive correlation between the sales in different regions, such as North America, Europe, and Rest of World. The data suggests that developing games for popular platforms like PS2, Wii, and X360 and releasing games in the Sports and Action genres could be a profitable decision for game developers. However, the correlation between other factors, such as year, rank, and review score, is not very strong. In conclusion, it is important for game developers to consider a combination of factors when developing games, such as platform popularity, genre popularity, and regional sales trends, in order to increase their chances of success in the competitive video game industry.