

Data Science Using Powerbi and Python

TITLE :-

Video Game Sales Data Analysis And Interactive Dashboard

Prepared By :- Vivek Yadav (BT23CSA035)

Ojaswa Awasthi (BT23CSA060)

Mangesh Thale (BT23CSA065)

Under the Supervision of

Mr. Santosh Kumar Sahu

Problem description :-

The video game industry is a dynamic and competitive market where understanding sales trends, customer preferences, and regional performance is crucial for publishers, developers, and marketers. However, the vast amount of sales data often remains underutilized, leaving key insights hidden.

This project aims to:

1. Conduct an exploratory data analysis (EDA) of the **Video Game Sales Dataset** to uncover trends, patterns, and outliers.
2. Visualize key metrics like total sales, genre popularity, and regional performance using intuitive and interactive dashboards.
3. Enable stakeholders to make data-driven decisions regarding game development, marketing, and regional targeting.

Key Questions to Address:

- Which genres and platforms are the most popular globally?
- How do sales trends differ across regions (e.g., North America, Europe, Japan)?
- What factors contribute to a game's commercial success (e.g., publisher, year of release)?
- Are there any significant shifts in gaming trends over time?

Application :-

- **Understanding Market Trends**

1. **Helps Identify which game genres and platforms are trending globally or regionally.**
2. **game developers focus on high-demand genres and optimize platform-specific releases.**

- **Strategic Decision-Making**

1. **Insights into regional preferences allow publishers to tailor marketing strategies for different markets.**
2. **Example: Prioritizing action games in North America while focusing on RPGs in Japan.**

- **Revenue Optimization**

1. **Analyze sales patterns to predict the best times for game launches.**
2. **Use historical data to price games more competitively based on platform and region performance.**

Application :-

- **Competitor Analysis**

1. **Monitor how competitors are performing in terms of sales by genre, platform, or region.**
2. **Helps identify gaps in the market and leverage them for new game development.**

- **Tracking Industry Shifts**

1. **Spot emerging trends, such as a rise in popularity of indie games or new platforms like VR/AR.**
2. **Provides early warnings for publishers to adapt to market changes.**

- **Dashboard as a Decision Support Tool**

1. **Interactive dashboards empower stakeholders to explore the data independently.**
2. **Real-time visualizations make it easier to present findings to non-technical audiences.**

Data Description :

The dataset contains information about video game sales globally, across various platforms, and over different time periods.

It includes data points such as sales figures, game details, and publisher information, offering a comprehensive view of the industry.

Key Attributes :-

- **Name:** Title of the video game.
- **Platform:** The gaming console/platform for which the game was released (e.g., PS4, Xbox, PC).
- **Year of Release:** The release year of the game.
- **Genre:** The type of game (e.g., Action, Sports, RPG).
- **Publisher:** Company responsible for publishing the game.
- **Regional Sales:** Sales data segmented by regions:
 - **NA_Sales:** North America sales (in millions).
 - **EU_Sales:** Europe sales (in millions).
 - **JP_Sales:** Japan sales (in millions).
 - **Other_Sales:** Sales in the rest of the world (in millions).
 - **Global_Sales:** Total worldwide sales (in millions).

Basic Data Wrangling/Munging/Prepressing steps :-

• Loading the Data

1. Imported the dataset into the analysis environment (e.g., Python).
2. Checked the data structure and previewed a few rows to understand its format.

• Handling Missing Values

1. Identified missing values in key columns (e.g.,Critic_score, User_Score,Critic_Counts,User_Counts).
2. Applied appropriate strategies:
3. Removed rows with extensive missing data.
4. Imputed missing values where possible (e.g., using the mean for Numeric Columns).

• Data Cleaning

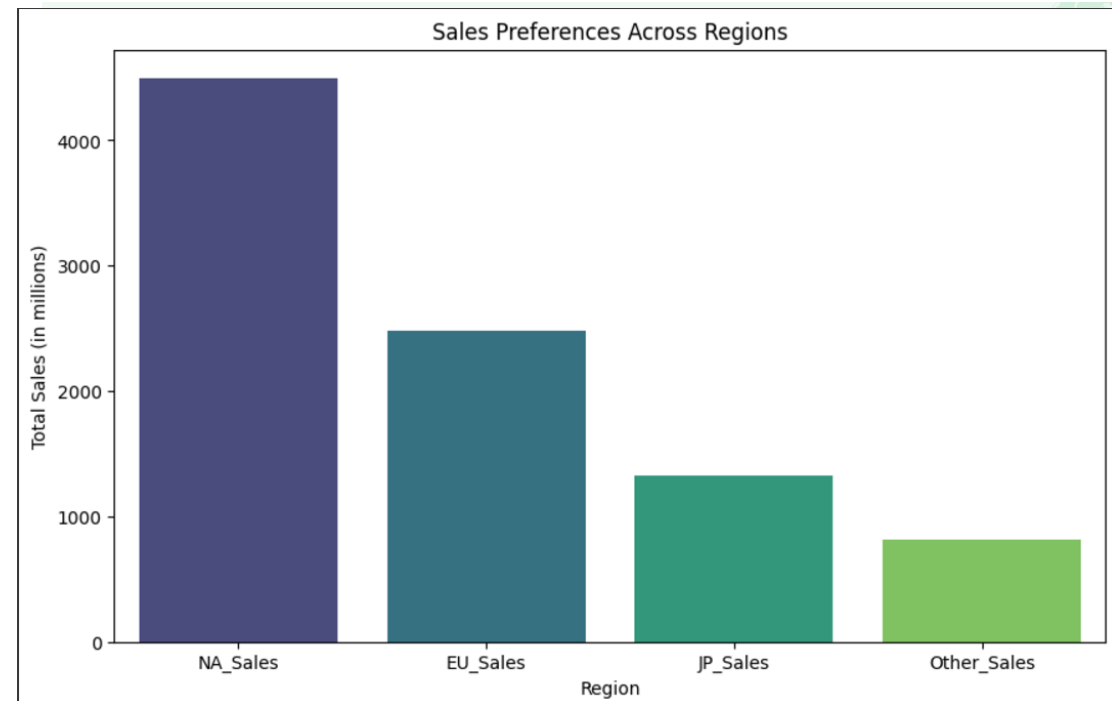
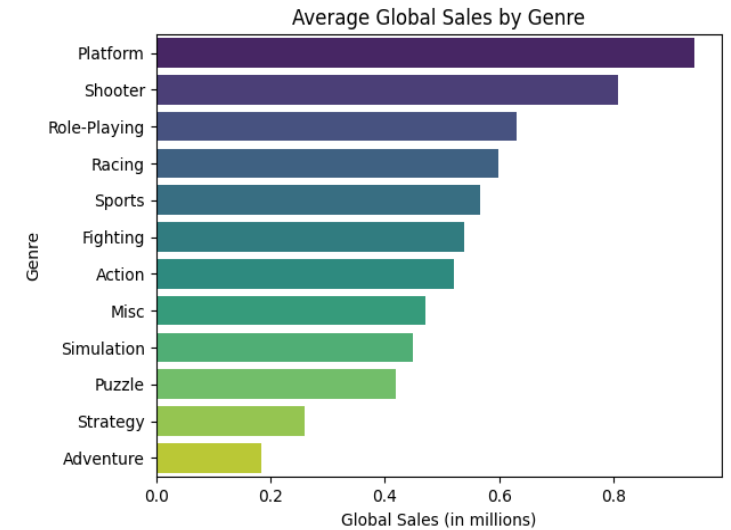
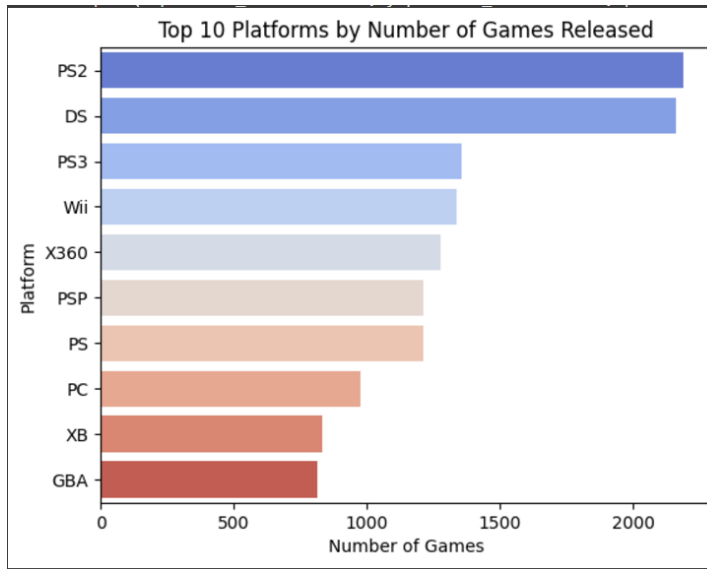
1. Removed duplicates to ensure data integrity.
2. Corrected obvious errors or anomalies in the data (e.g., invalid years).

• Feature Engineering

1. Created new columns for analysis:
2. Combined regional sales to compute total global sales if not already present.
3. Added a column for decade-based grouping of release years.

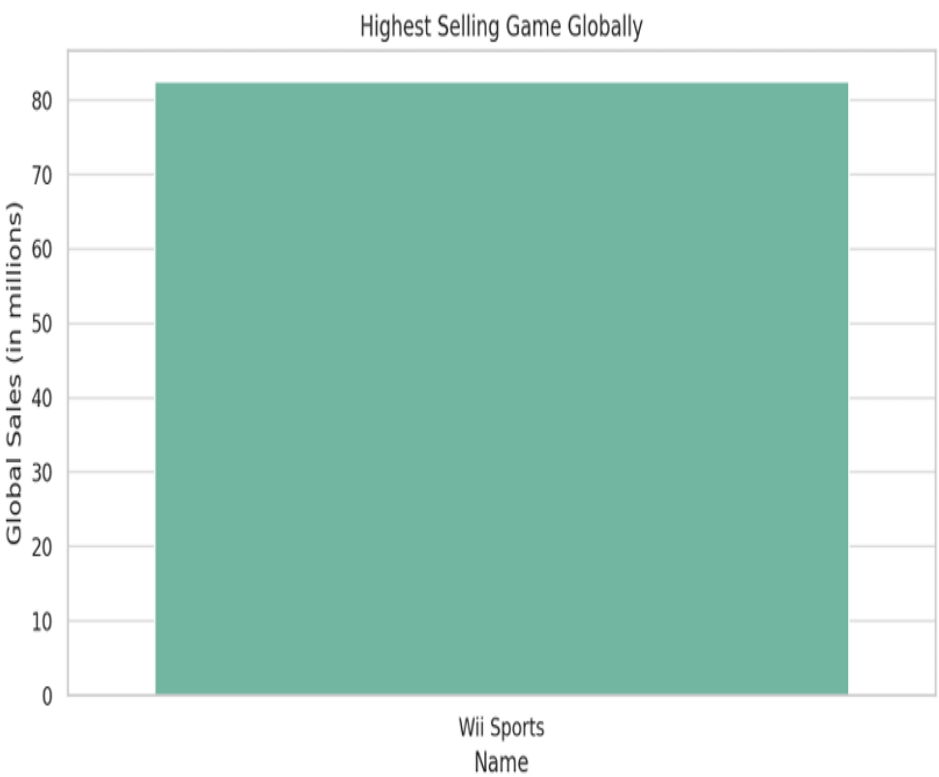
Video games sales

Exploratory Data Analysis (EDA)

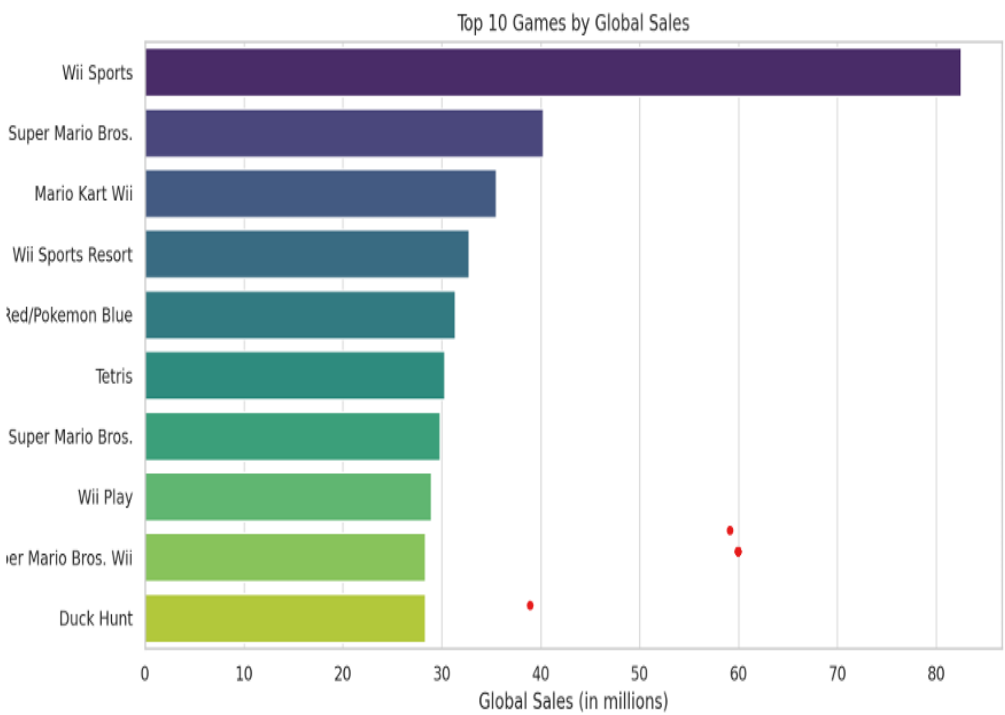


EDA Top Questions:-

Q1: 1. Which game has the highest global sales?

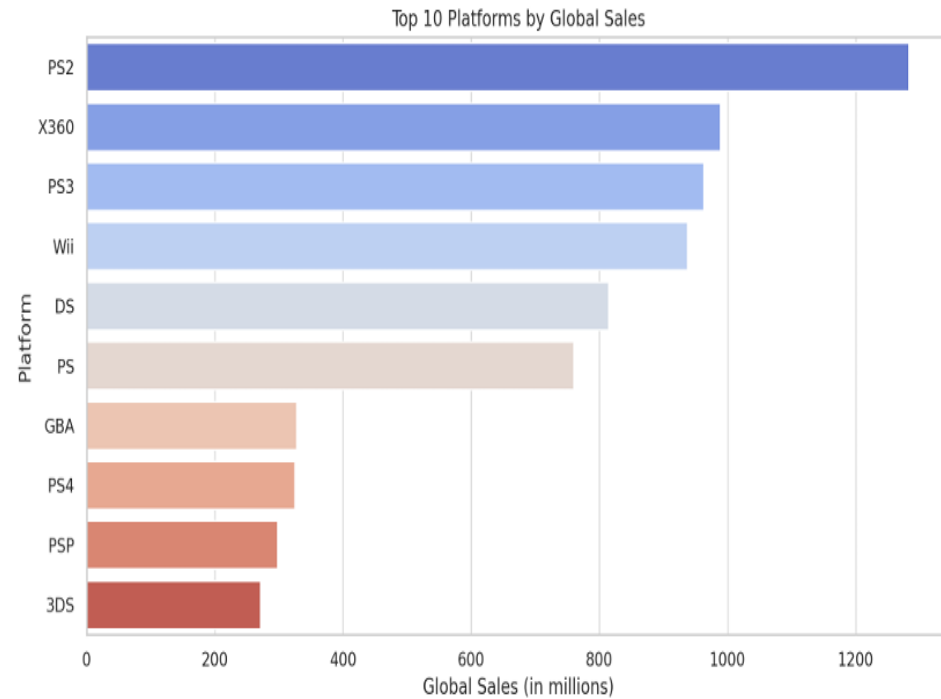


Q2: 2. Top 10 games by global sales.

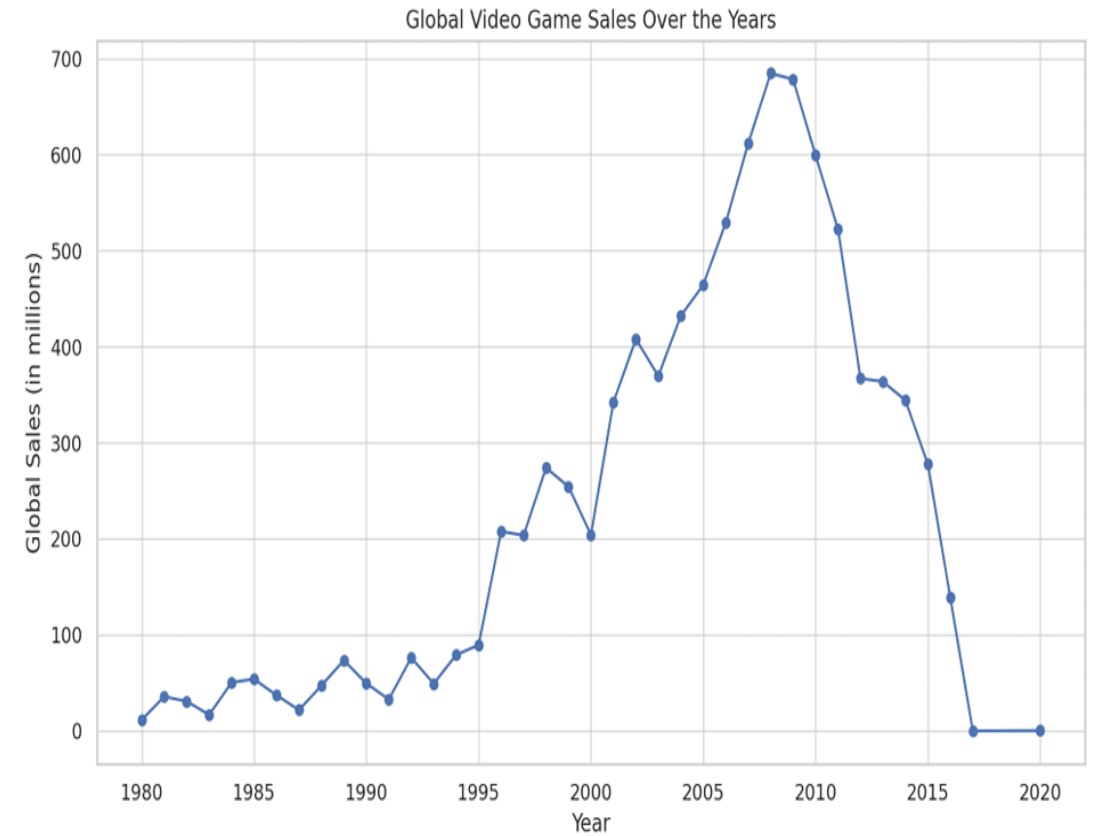


EDA Top Questions:-

Q3: 3. Platform with most global sales.

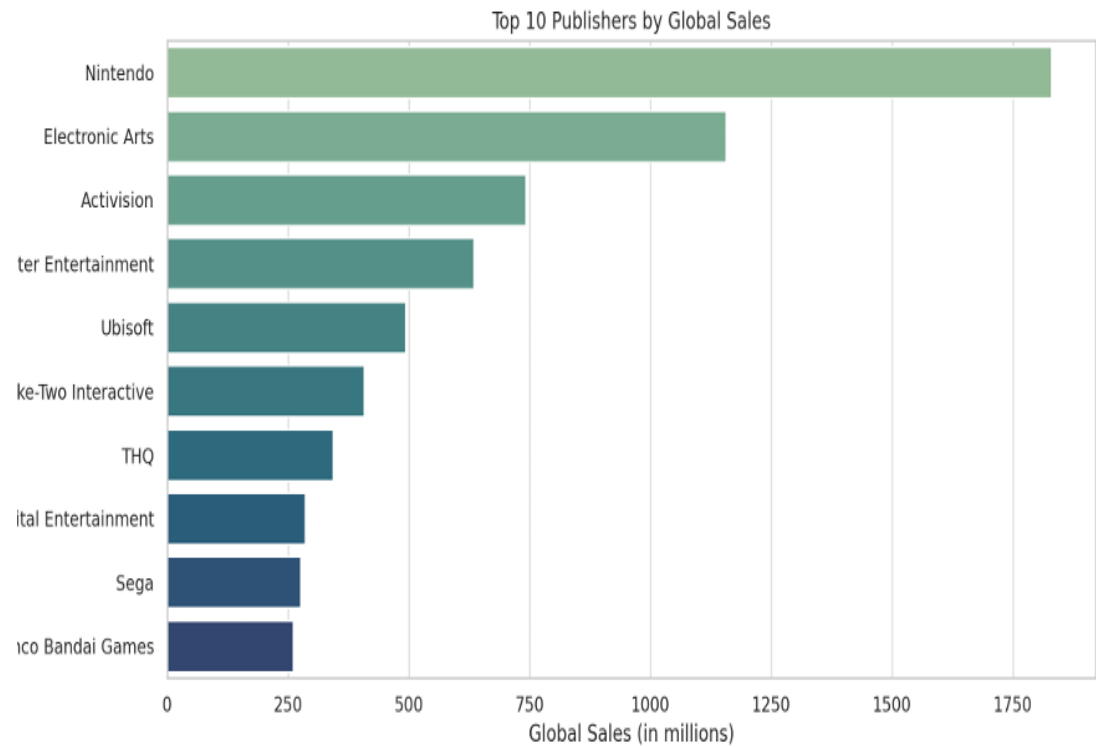


Q4: 4. Global sales trend over years.

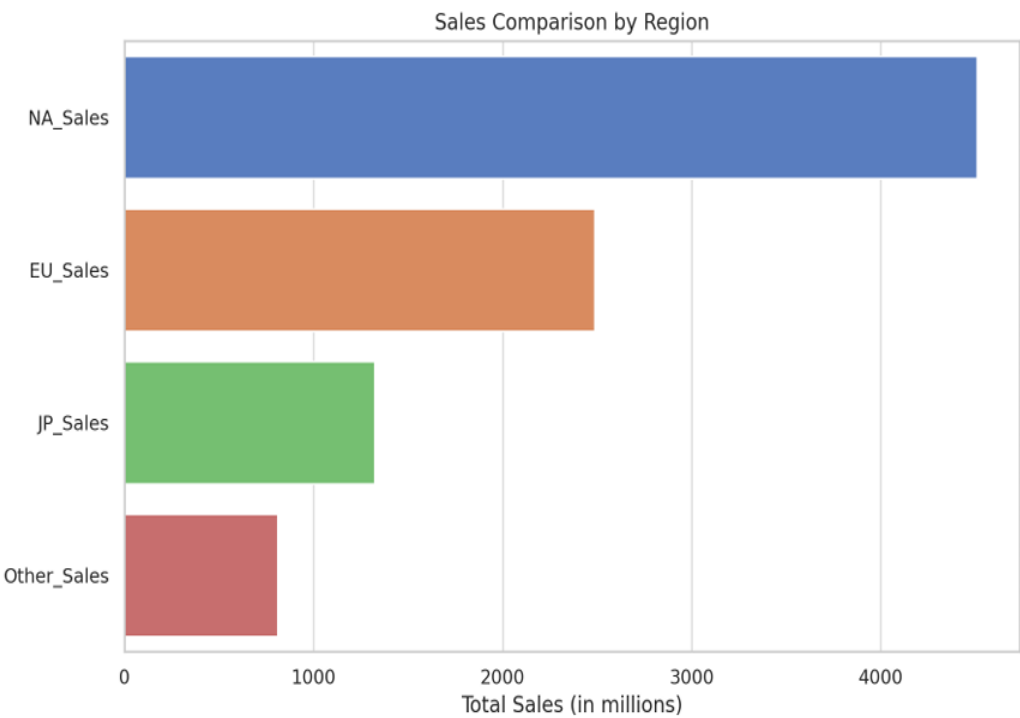


EDA Top Questions:-

Q5: 5. Publisher with highest global sales.

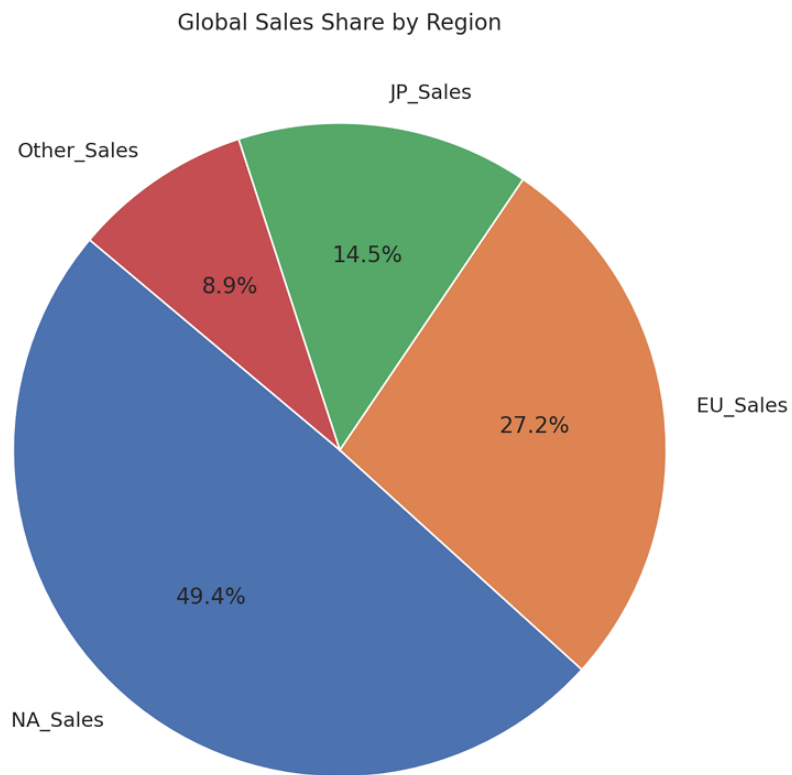


Q6: 6. Regional sales comparison (NA, EU, JP, Others).

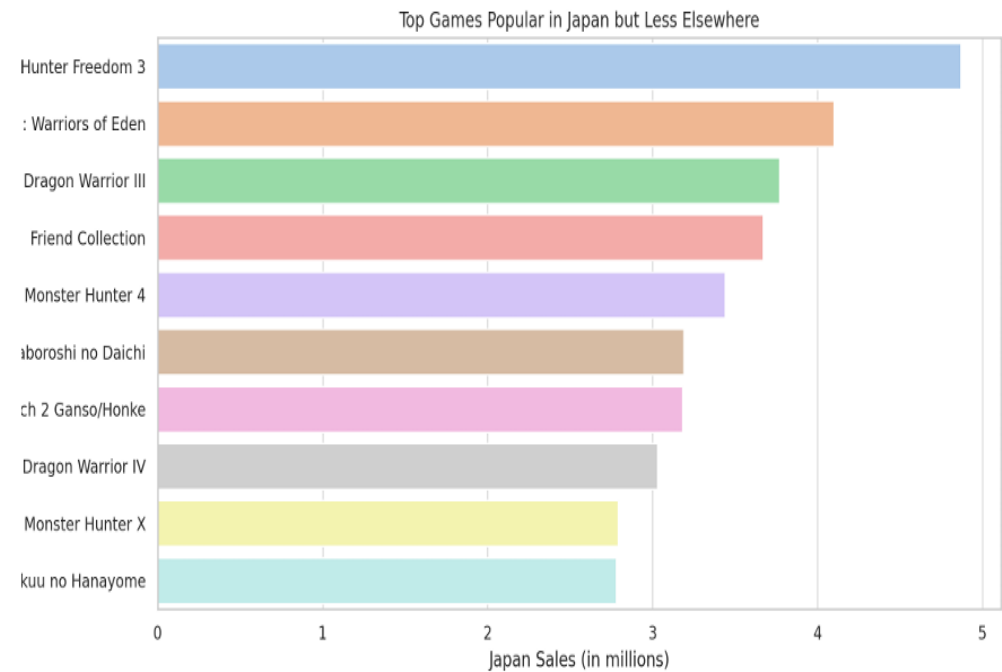


EDA Top Questions:-

Q7: 7. Region with highest global sales share.

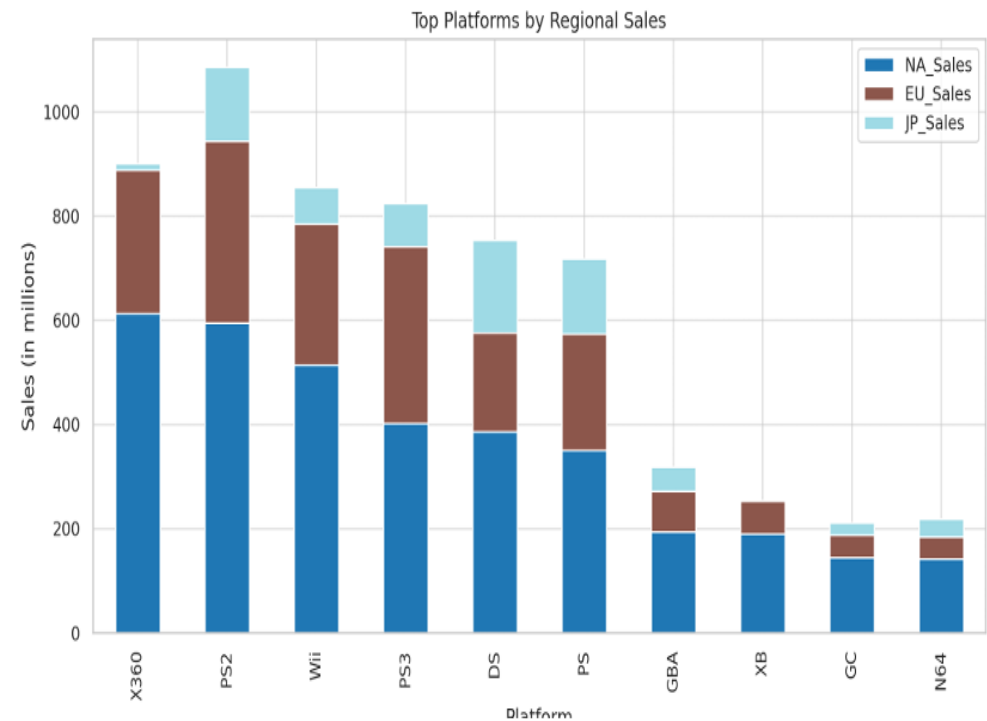


Q8: 8. Games successful in Japan but not elsewhere.

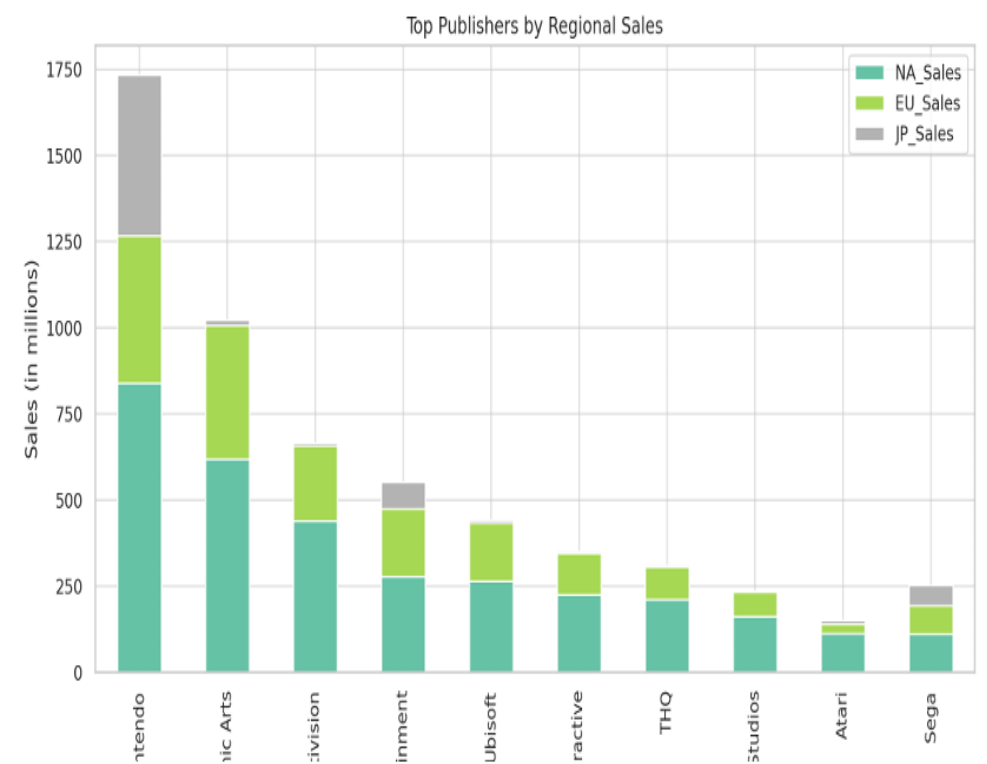


EDA Top Questions:-

Q9: 9. Top platform in each region.

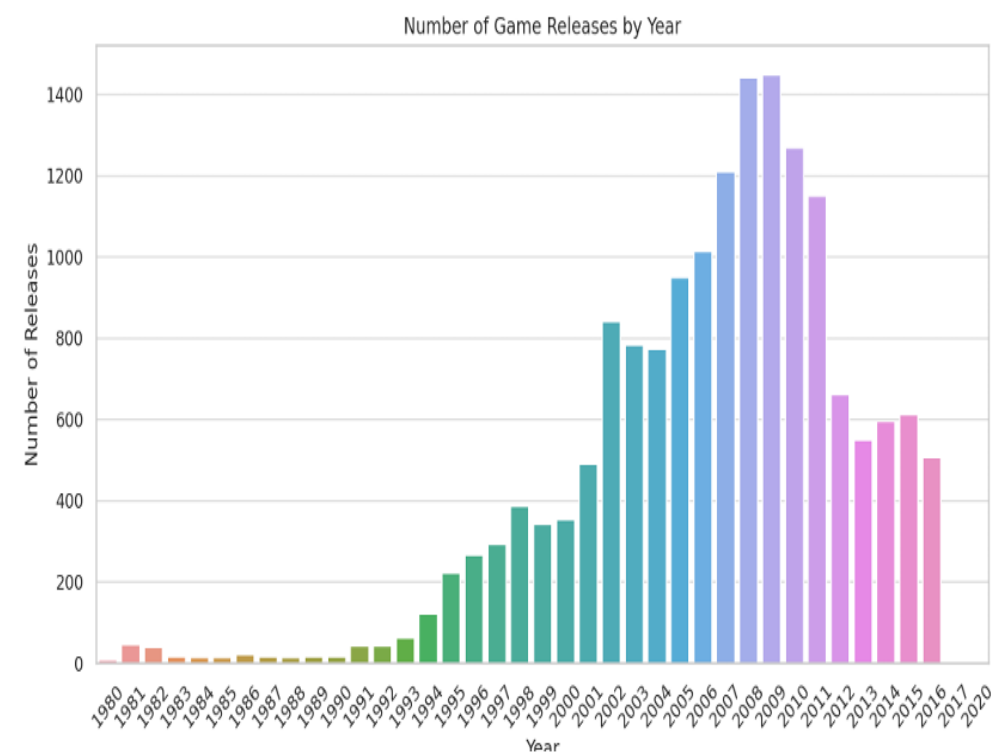


Q10: 10. Top publisher in each region.

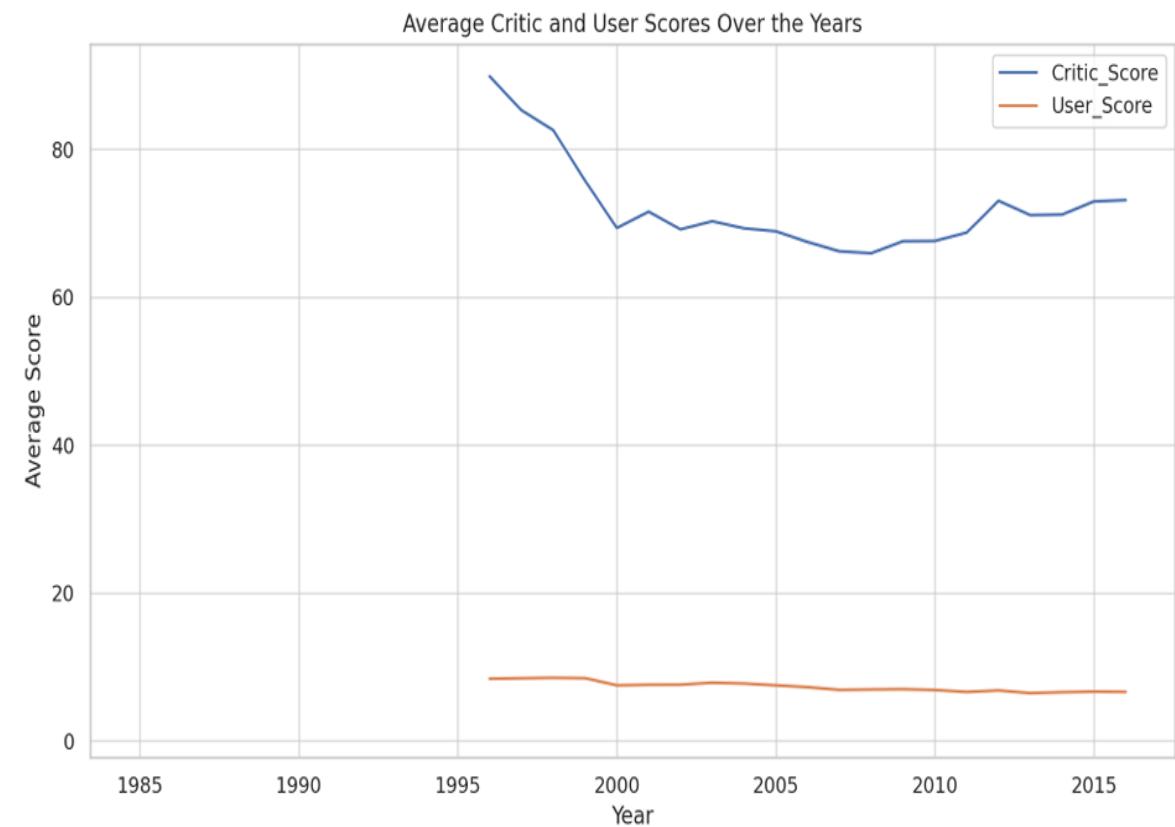


EDA Top Questions:-

Q11: 11. Year with most game releases.



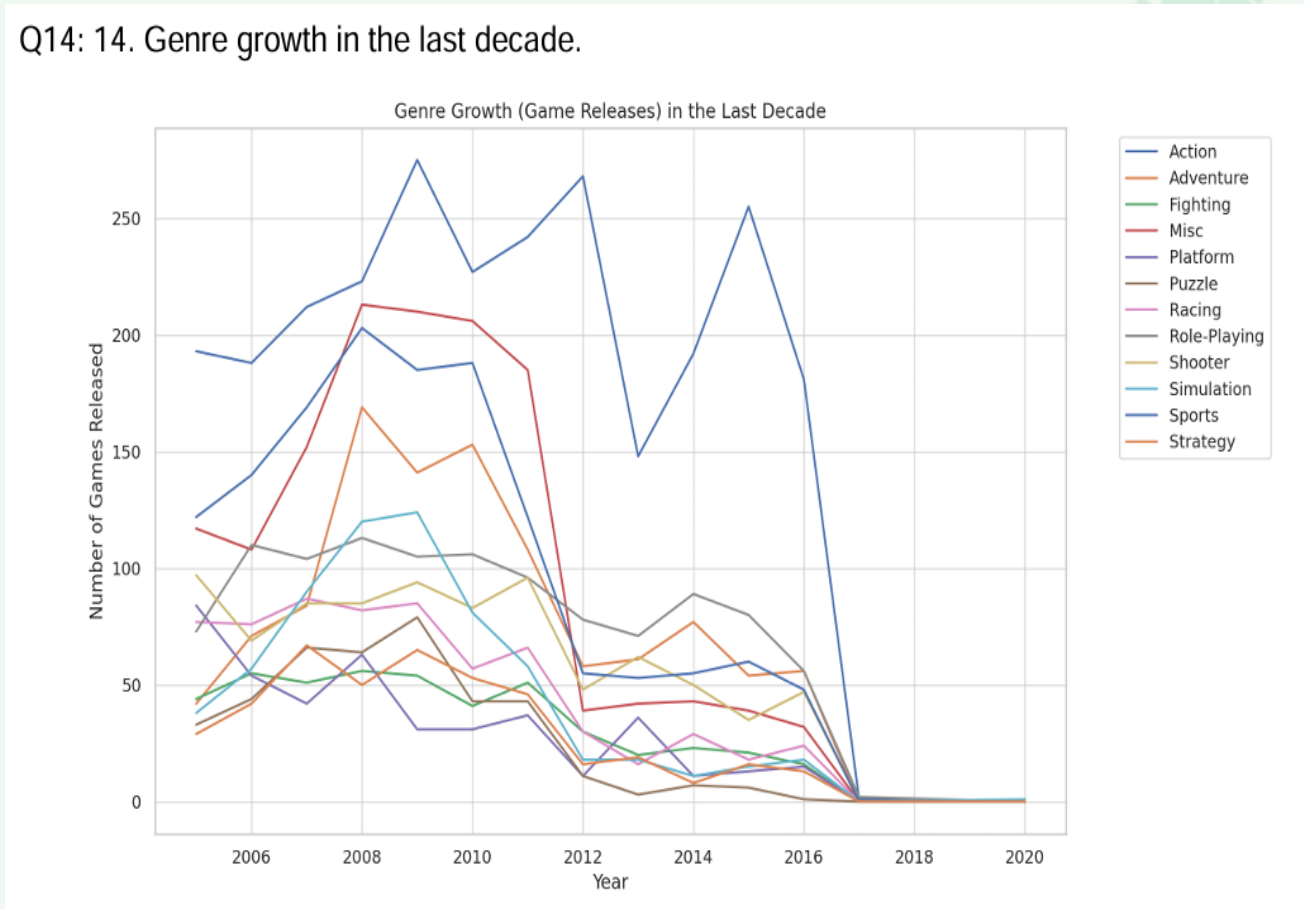
Q12: 12. Trends in critic and user scores over years.



Q13: 13. Change in average global sales per game.

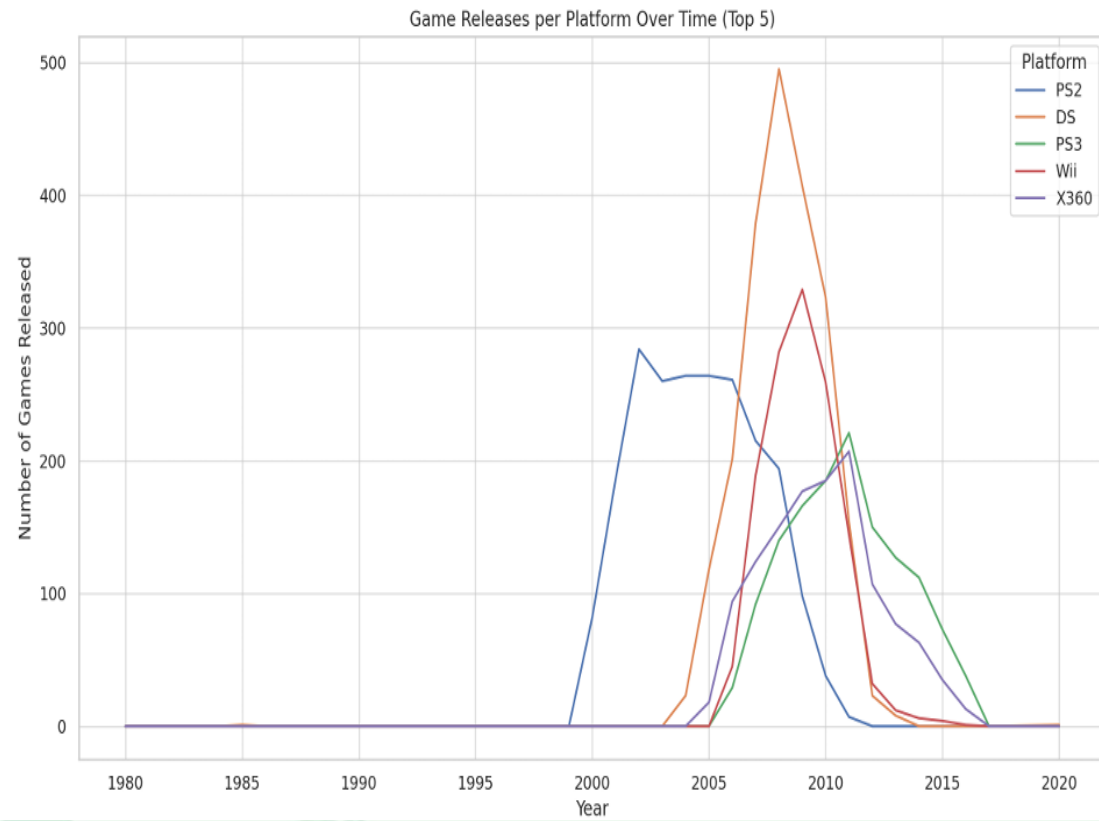
The line graph illustrates the average global sales per game in millions from 1980 to 2020. The y-axis represents 'Average Sales (millions)' ranging from 0 to 4, and the x-axis represents 'Year' from 1980 to 2020. The sales show a significant peak around 1989, followed by a sharp decline and then a period of relative stability with minor fluctuations until 2017, after which there is a slight upward trend.

Year	Average Sales (millions)
1980	1.3
1981	0.8
1982	0.8
1983	1.0
1984	3.6
1985	3.9
1986	1.8
1987	1.4
1988	3.2
1989	4.3
1990	3.1
1991	0.8
1992	1.8
1993	0.8
1994	0.7
1995	0.4
1996	0.8
1997	0.7
1998	0.7
1999	0.7
2000	0.6
2001	0.7
2002	0.5
2003	0.5
2004	0.6
2005	0.5
2006	0.5
2007	0.5
2008	0.5
2009	0.5
2010	0.5
2011	0.5
2012	0.6
2013	0.7
2014	0.6
2015	0.5
2016	0.3
2017	0.1
2018	0.2
2019	0.3
2020	0.4

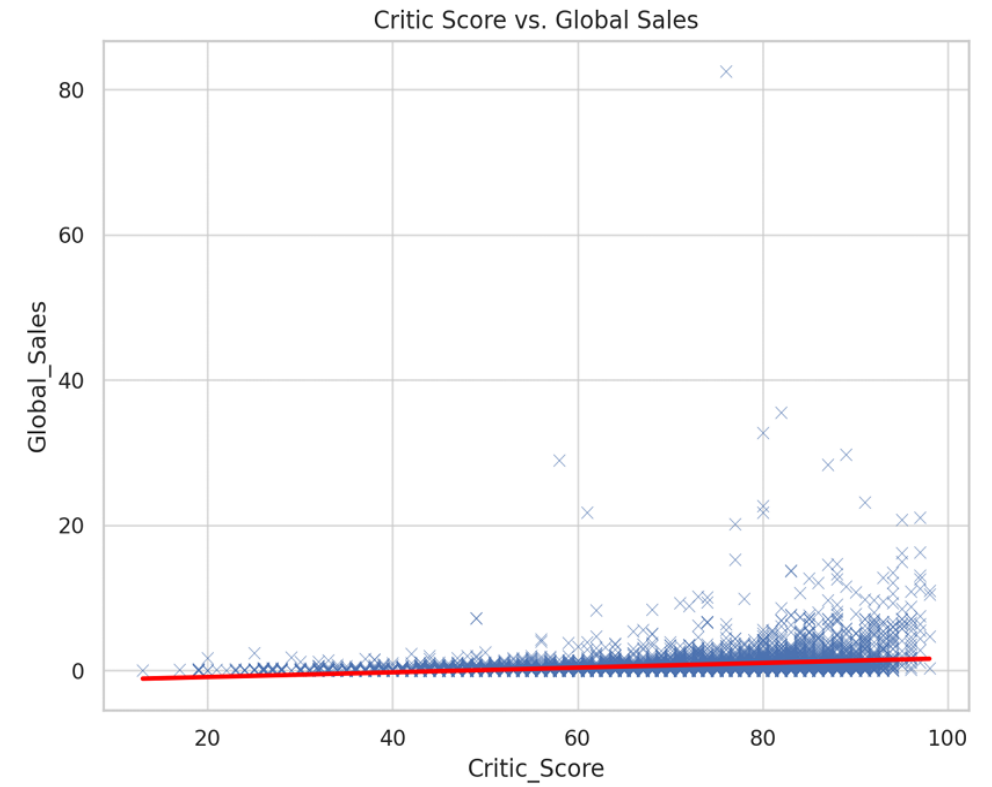


EDA Top Questions:-

Q15: 15. Game release trend per platform.

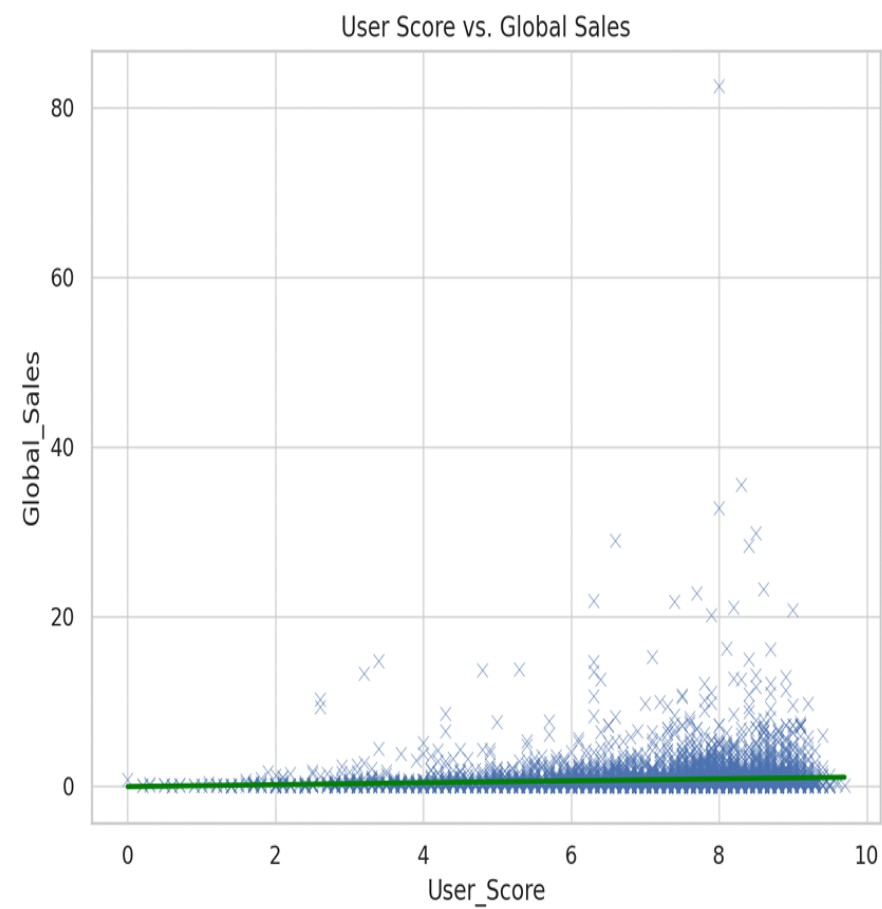


Q16: 16. Critic Score vs. Global Sales.

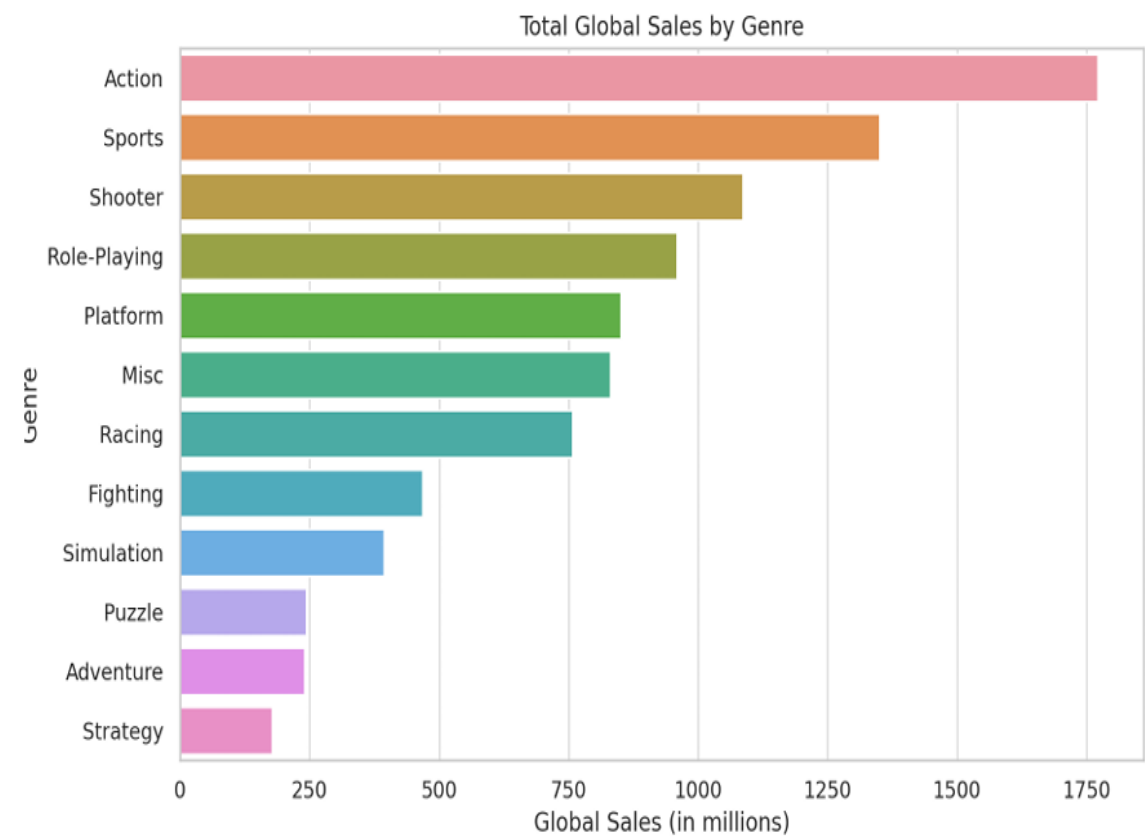


EDA Top Questions:-

Q17: 17. User Score vs. Global Sales.

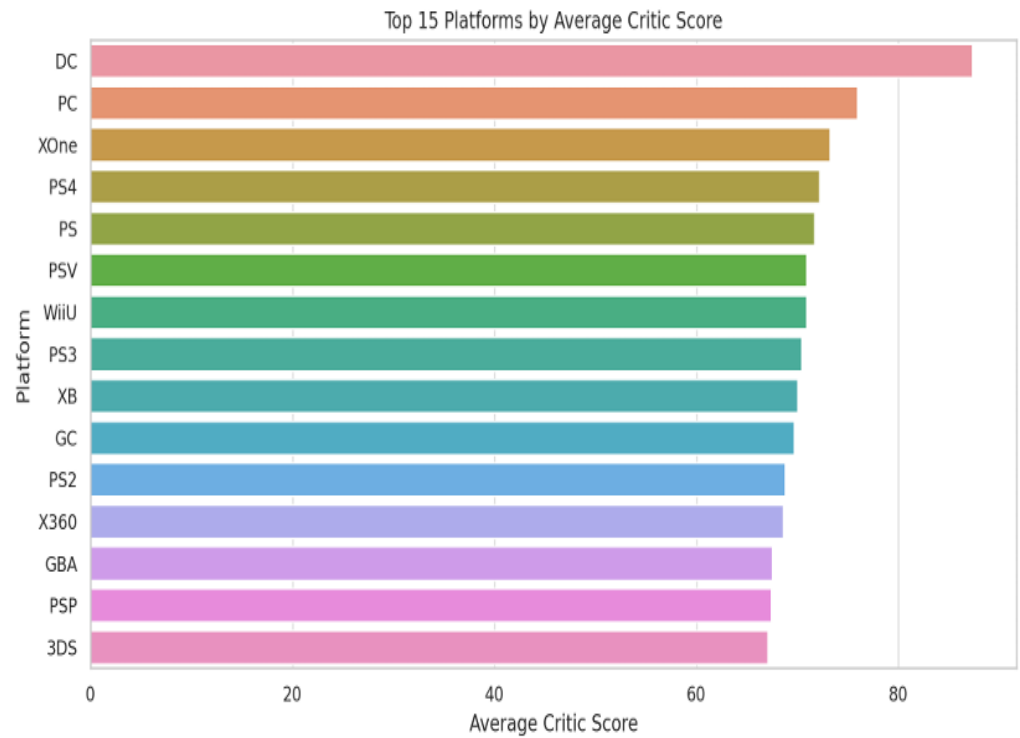


Q18: 18. Global Sales by Genre.

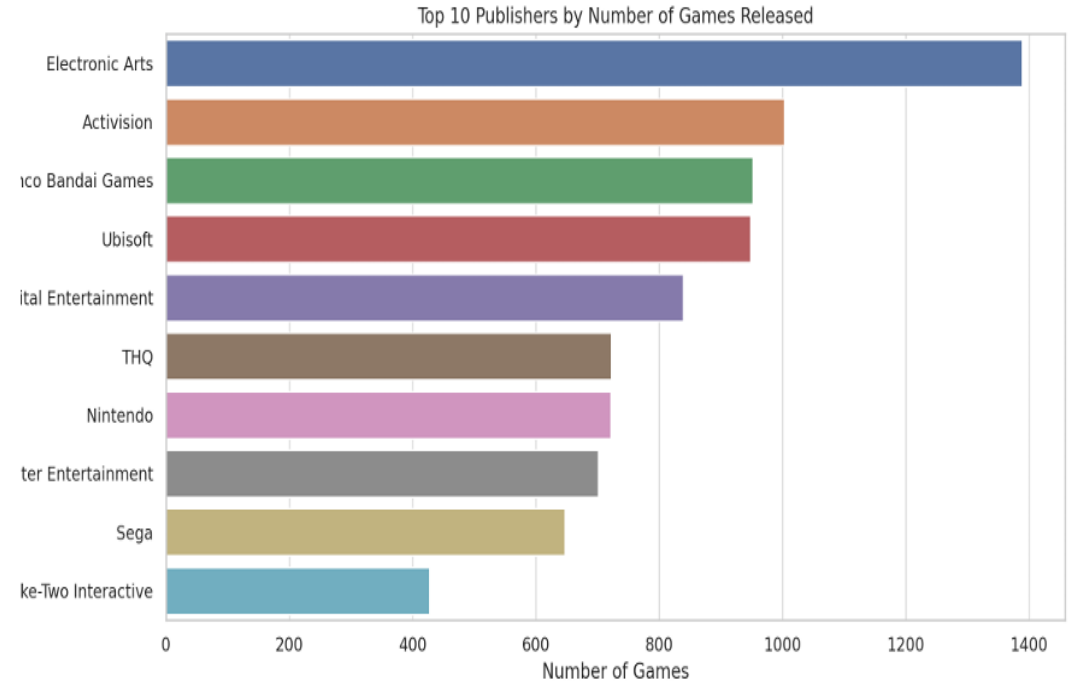


EDA Top Questions:-

Q19: 19. Top platforms by average Critic Score.

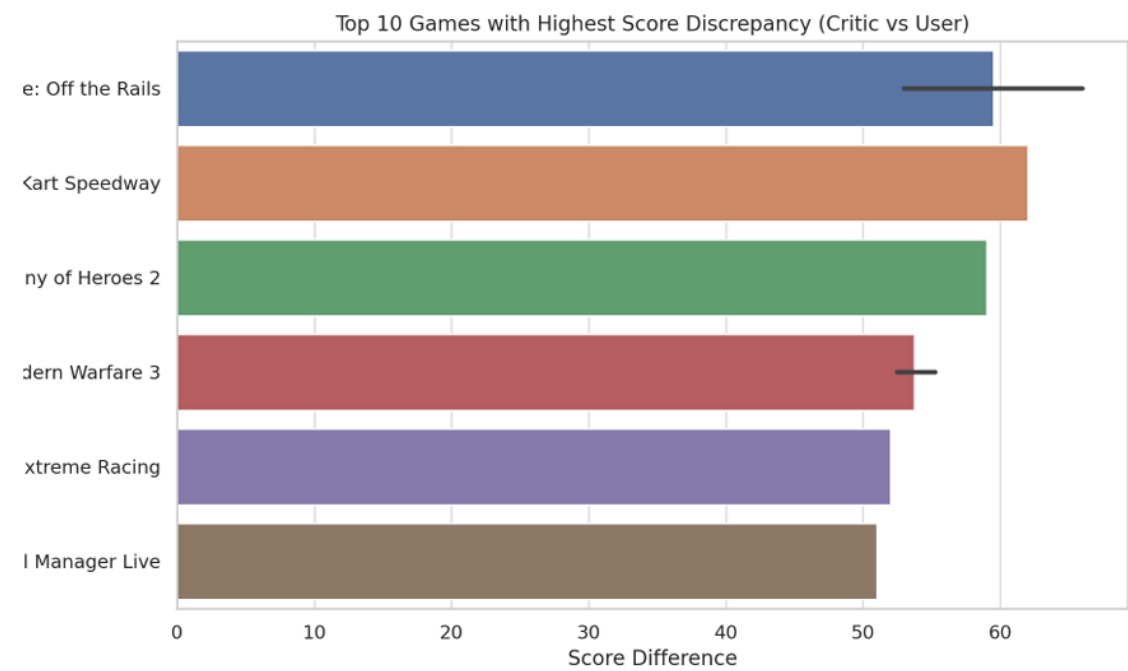


Q20: 20. Top publishers by game count.

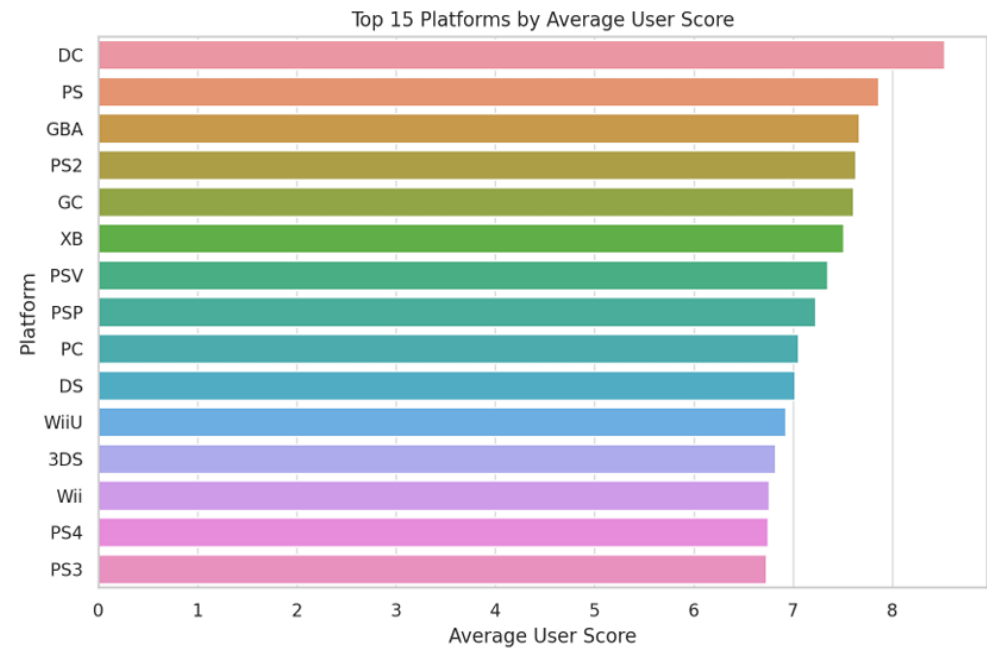


EDA Top Questions:-

Q21: 21. Games with biggest critic-user score gap.

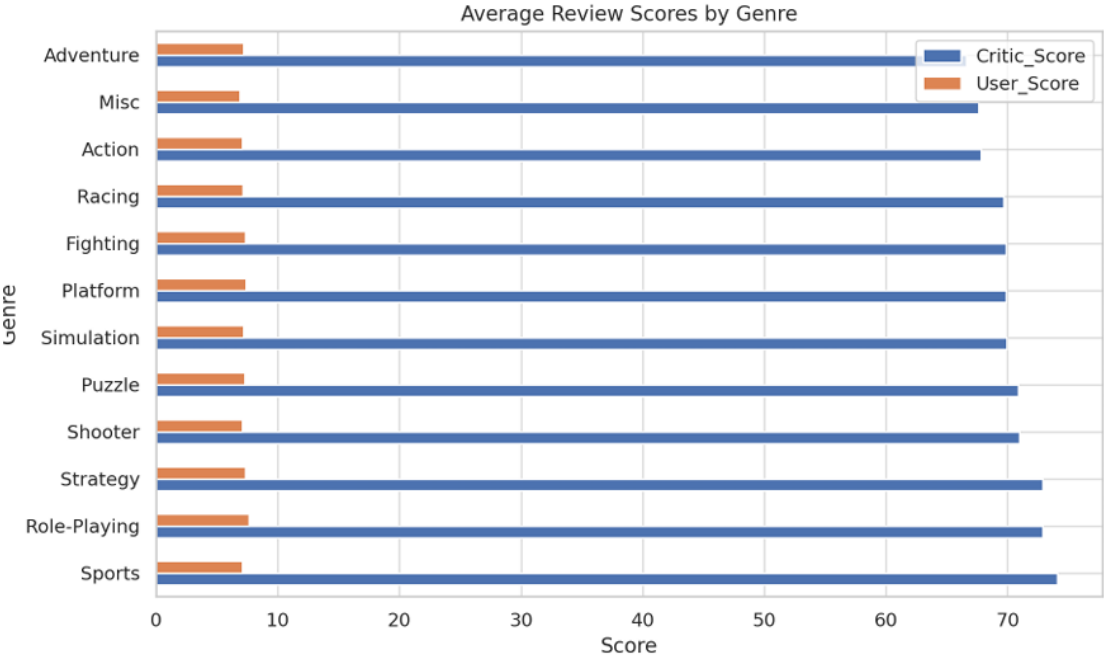


Q22: 22. Top platforms by user score.

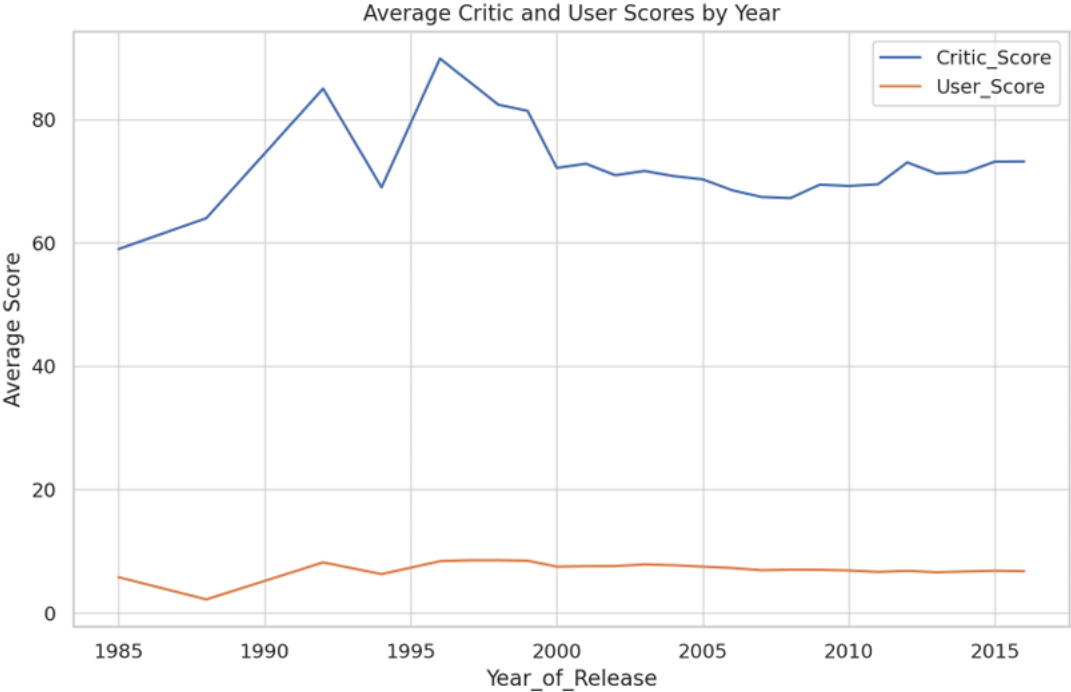


EDA Top Questions:-

Q23: 23. Average review scores by genre.

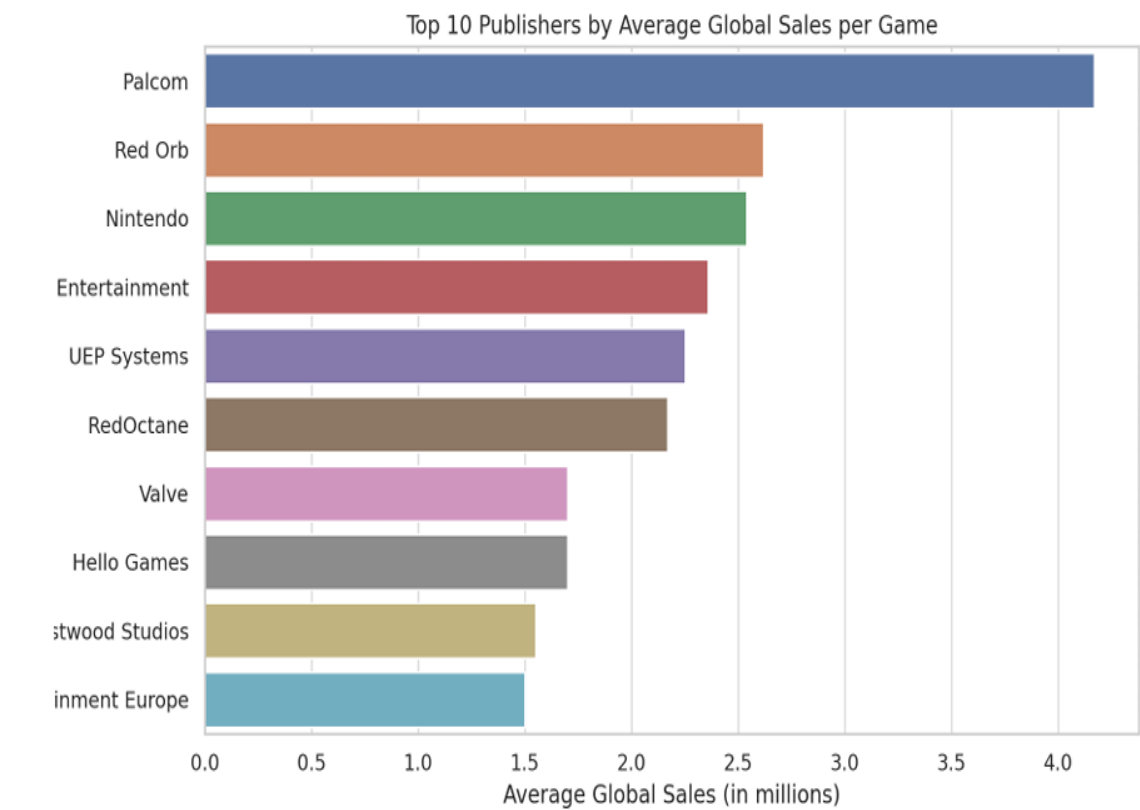


Q24: 24. Average review scores by year.

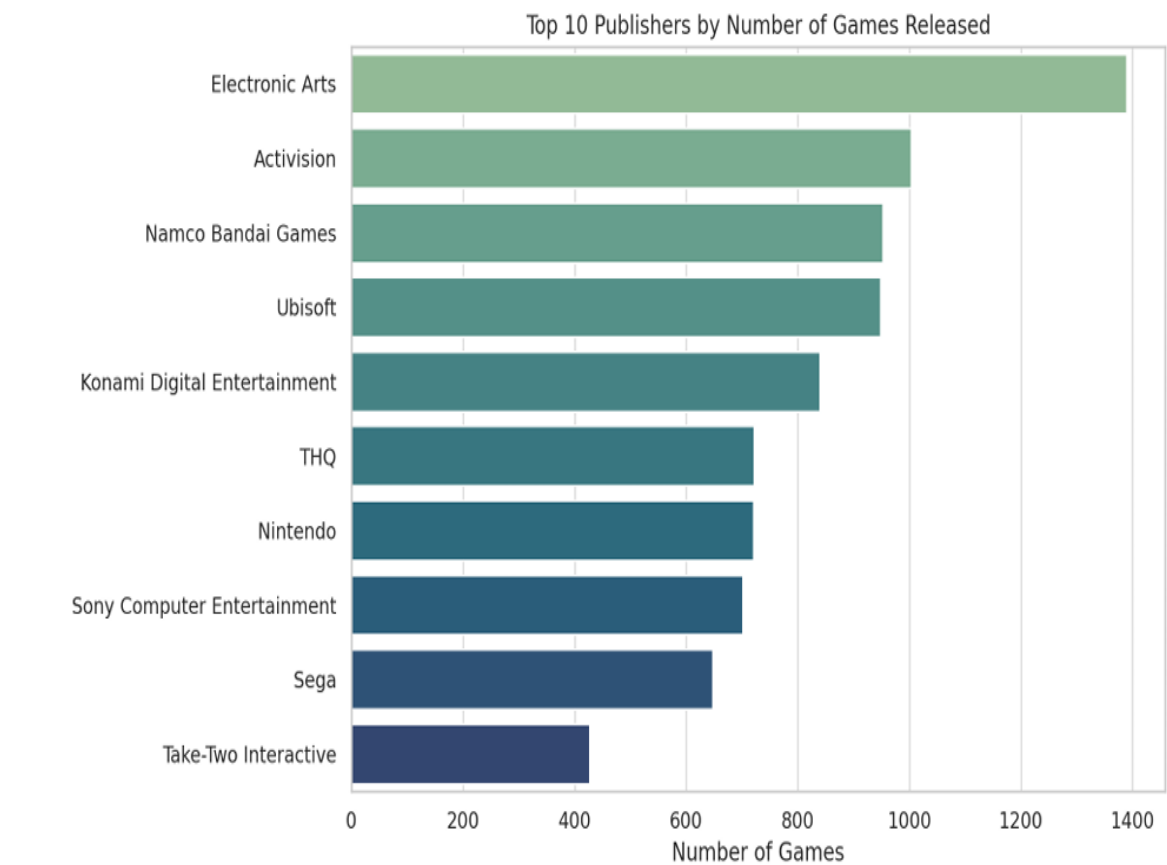


EDA Top Questions:-

Q25: 25. Top publishers by avg sales/game.

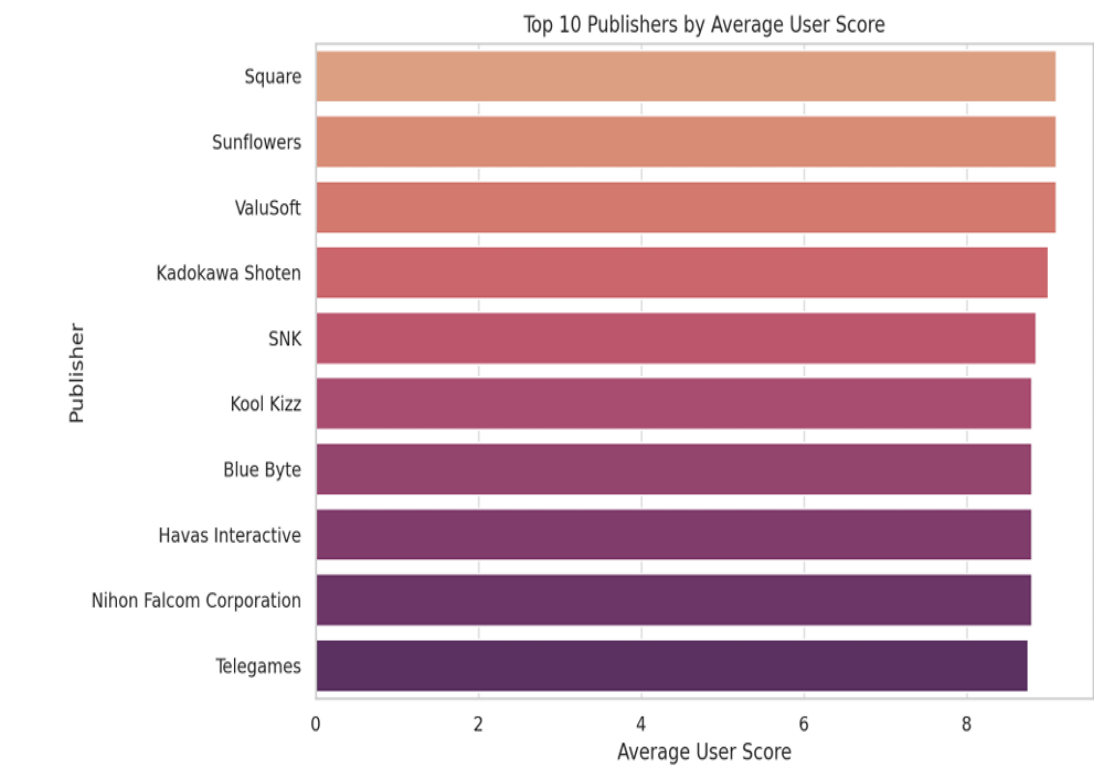


Q26: 26. Publisher with most games released.

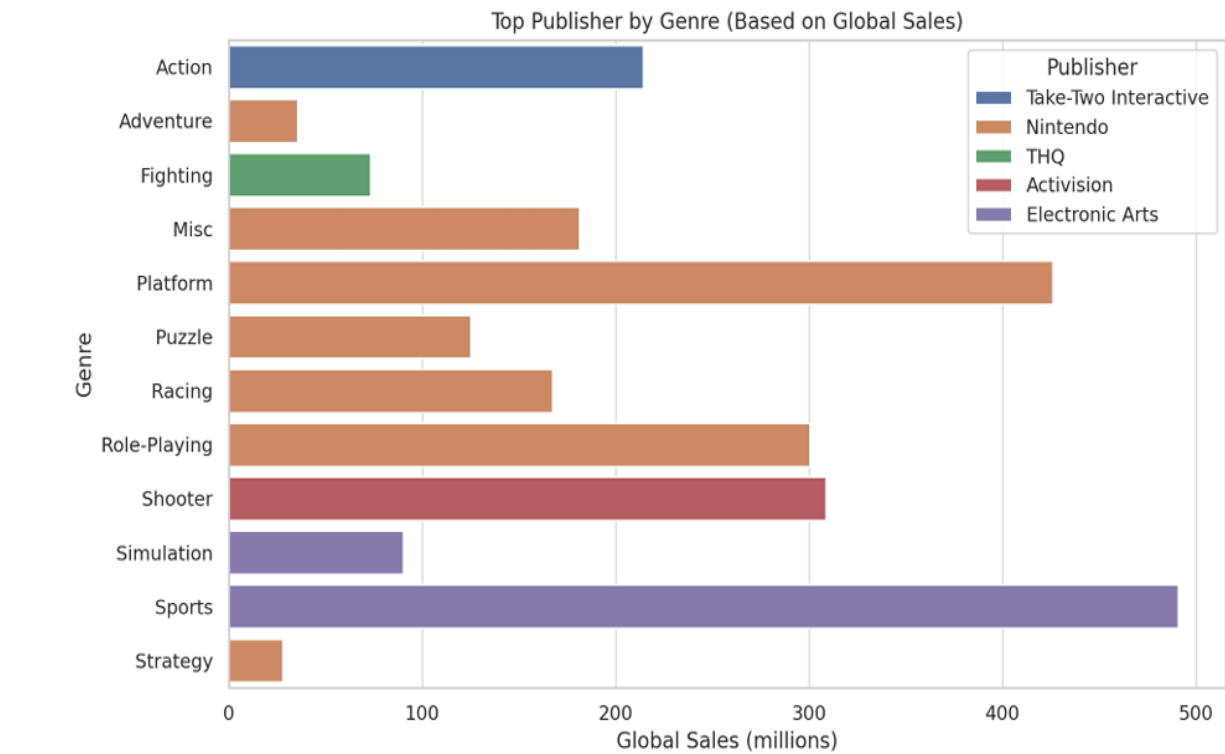


EDA Top Questions:-

Q27: 27. Publisher with highest avg user score.

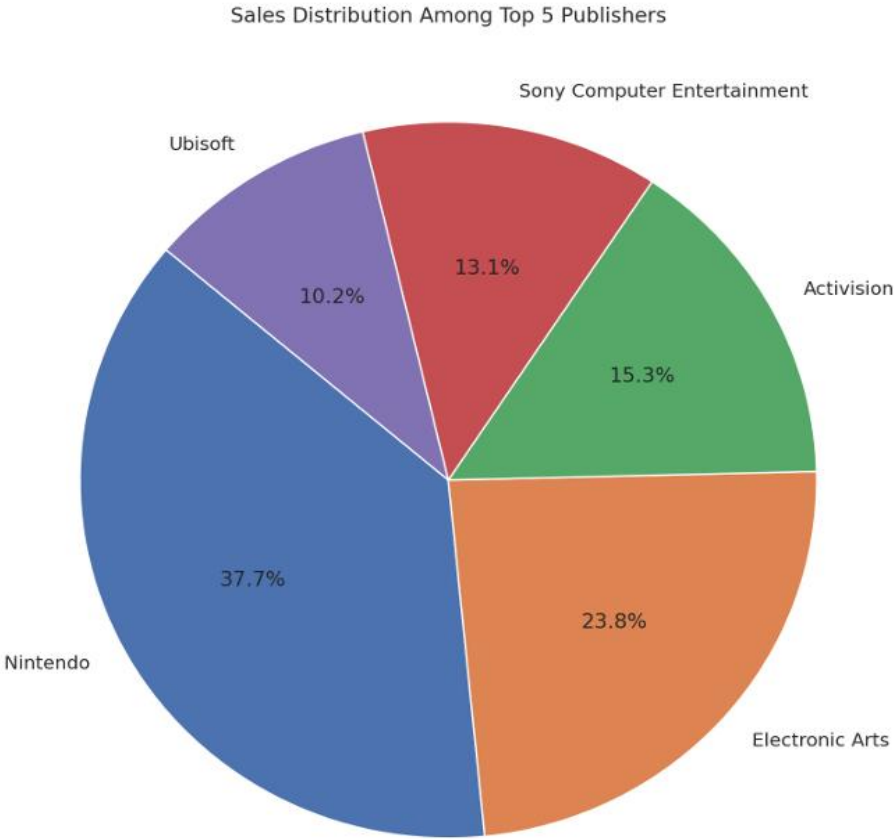


Q28: 28. Publisher dominance by genre.

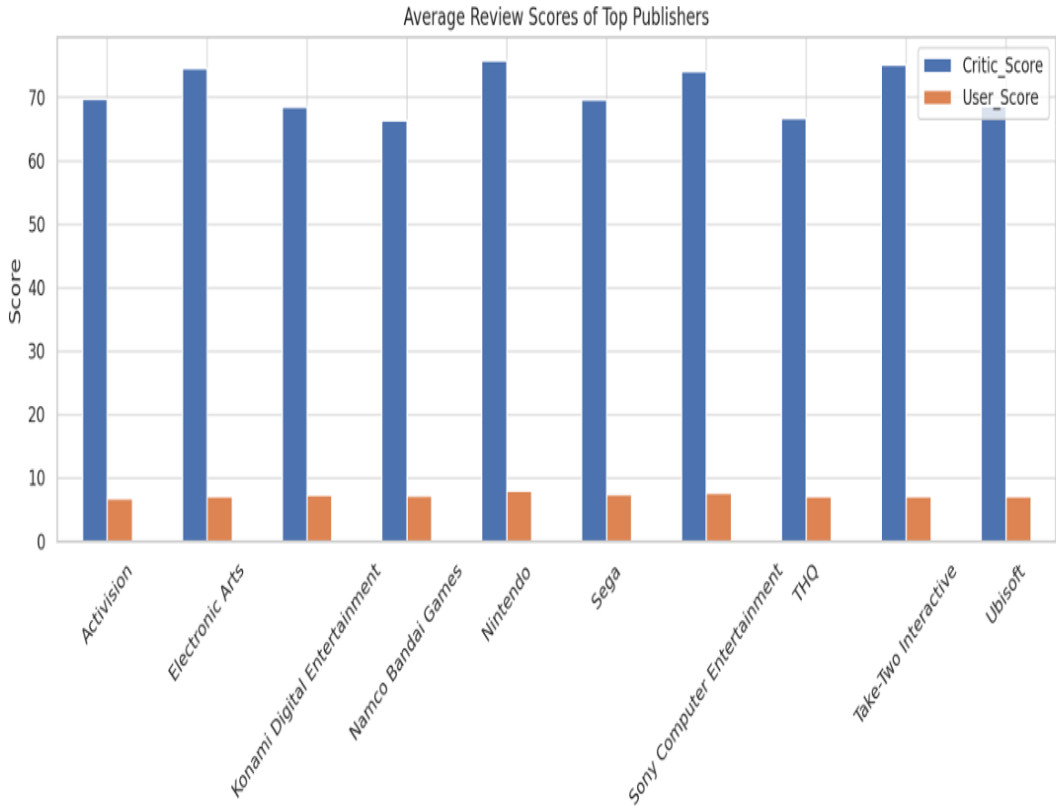


EDA Top Questions:-

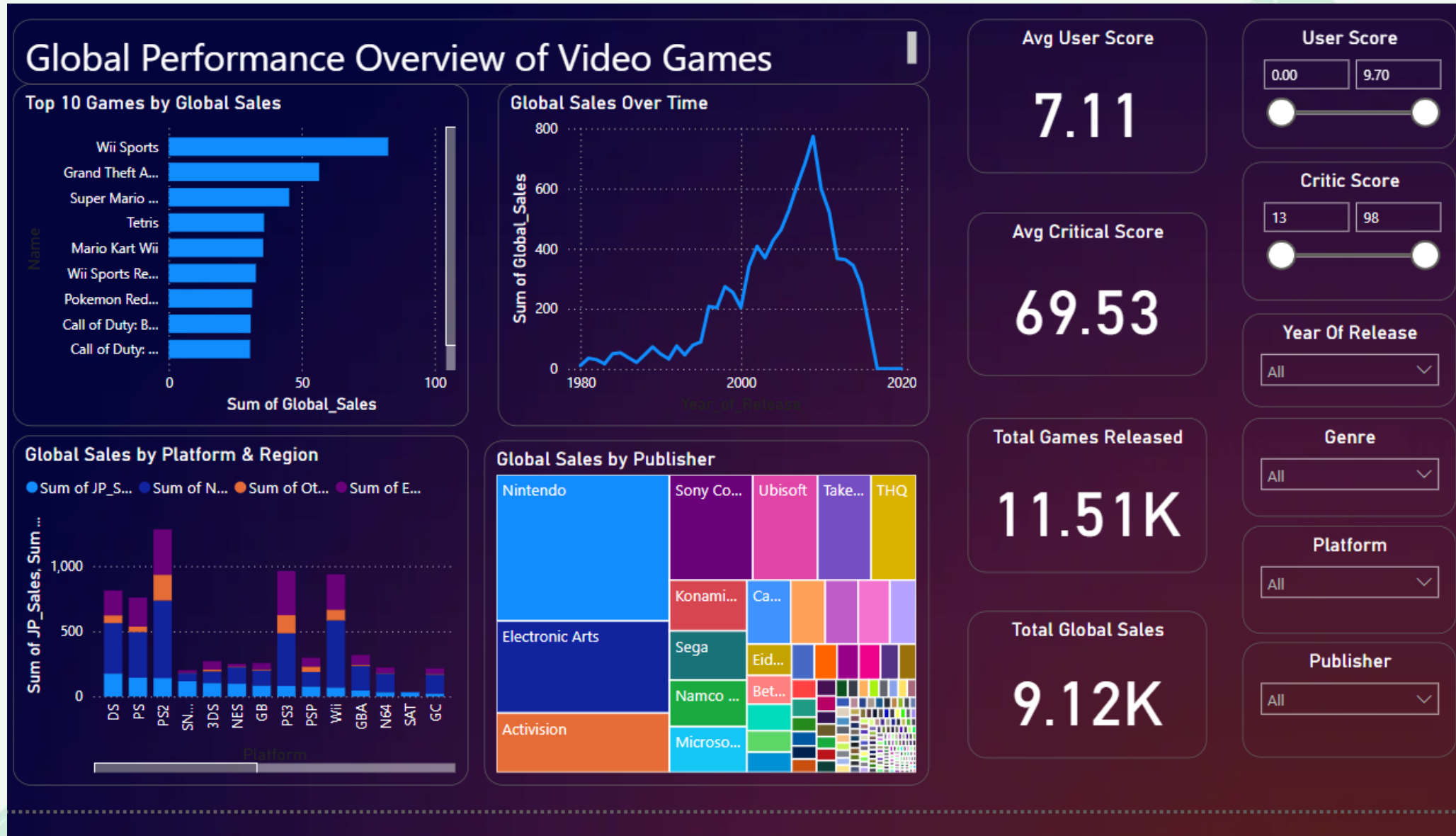
Q29: 29. Sales distribution among top 5 publishers.



Q30: 30. Publishers with best review consistency.



INTERACTIVE DYNAMIC DASHBOARD :-



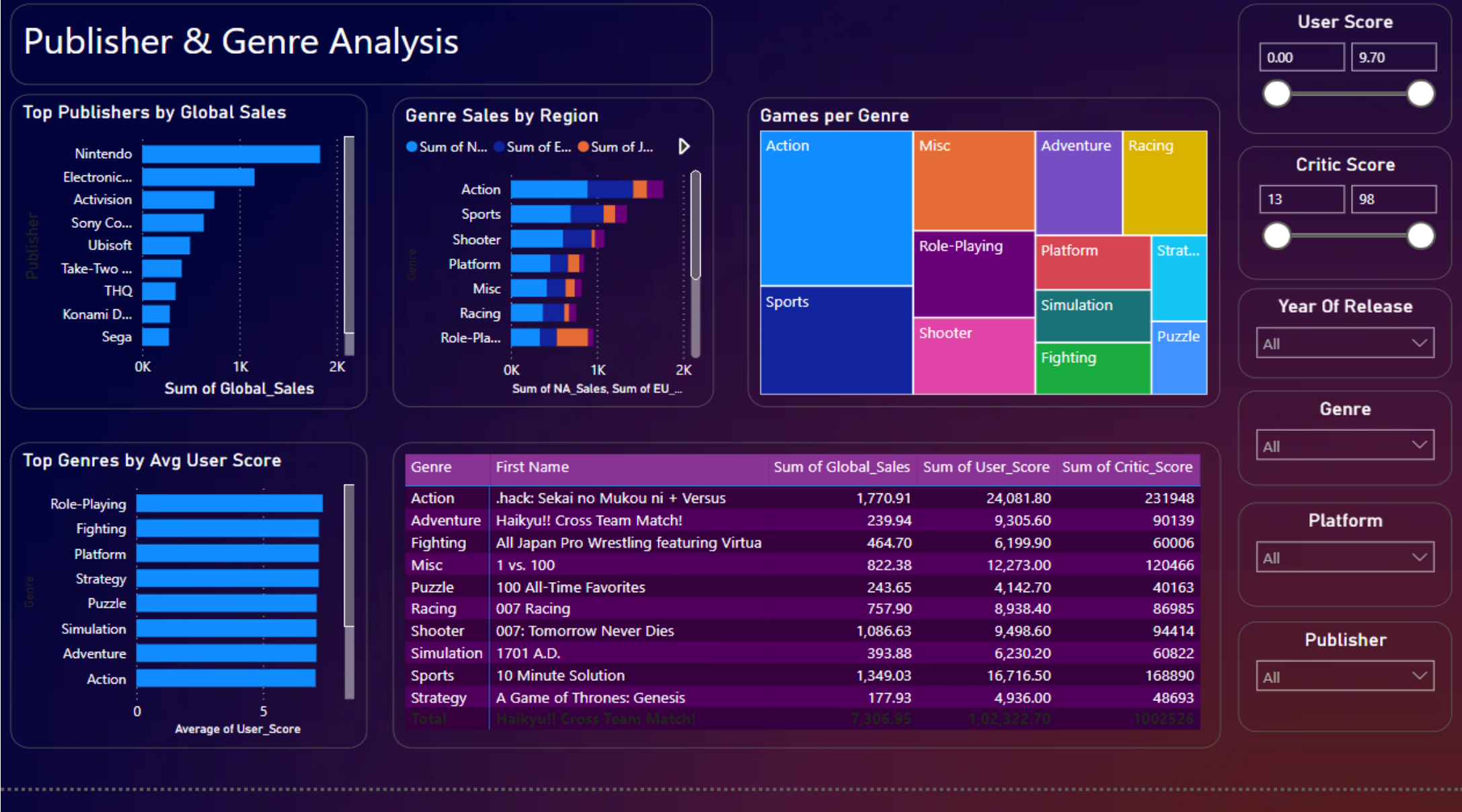
INTERACTIVE DYNAMIC DASHBOARD :-



INTERACTIVE DYNAMIC DASHBOARD :-



INTERACTIVE DYNAMIC DASHBOARD :-



Conclusion:-

- **Key Insights**

- I. **Market Trends:** Identified the most popular game genres and platforms globally and regionally.
- II. **Regional Preferences:** Observed significant differences in sales patterns across North America, Europe, and Japan.
- III. **Sales Drivers:** Highlighted factors influencing game success, such as genre, platform, and publisher performance.
- IV. **Temporal Trends:** Detected shifts in gaming trends over time, reflecting industry evolution.

- **Impact of EDA and Dashboard**

- I. EDA helped uncover hidden patterns and relationships in the dataset, providing a strong foundation for decision-making.
- II. The interactive dashboard allows stakeholders to explore data independently, fostering better insights and strategies.

- **Applications of Insights**

- I. Supports data-driven marketing strategies and game development decisions.
- II. Helps publishers allocate resources effectively by identifying high-performing genres and regions.
- III. Facilitates monitoring of industry trends and competitive landscapes.

- **Challenges and Future Scope**

- I. Challenges included handling missing data and addressing data inconsistencies.
- II. Future improvements could involve incorporating more recent data, additional metrics like user reviews, or integrating external datasets (e.g., market trends or competitor analysis).

- **Final Takeaway**

This project demonstrates the power of data in transforming raw information into actionable insights, ultimately driving better business decisions in the video game industry.

References:-

- **Dataset Source:** [https://www.kaggle.com/datasets/sidtwr/videogames-sales-dataset?select=Video Games Sales as at 22 Dec 2016.csv](https://www.kaggle.com/datasets/sidtwr/videogames-sales-dataset?select=Video+Games+Sales+as+at+22+Dec+2016.csv)
- **Tools Used:** Python (pandas, numpy, matplotlib, seaborn), Power BI.

Thank You