Exploratory Data Analysis of Video Game Sales Dataset

1. Introduction

This report presents an exploratory data analysis (EDA) of the Video Game Sales dataset. The objective is to understand the distribution and trends in video game releases and sales across different platforms, genres, regions, and years.

Tools Used:
Python, Pandas, Matplotlib, Seaborn, Jupyter Notebook

Dataset:
vgsales.csv — includes video game metadata and sales in millions.

2. Data Overview

- Total entries: **16,598**
- Key columns:
  - Name, Platform, Year, Genre, Publisher
  - NA_Sales, EU_Sales, JP_Sales, Other_Sales, Global_Sales

3. Data Cleaning

- **Missing Values**:
  - Year: 271 missing
  - Publisher: 58 missing

- **Handling**:
  - Dropped rows with missing Year or Publisher
  - Converted Year to integer format

  Final cleaned shape: 16,291 rows × 11 columns

4. Trend Analysis

🎮 Top 5 Genres Over Time

A line plot of the number of games released per genre over time shows:

- **Action** and **Sports** consistently dominate releases.
- **Shooter**, **Role-Playing**, and **Platform** games are also popular, with some fluctuation.

5. Sales Distribution by Genre

A boxplot of Global_Sales by Genre reveals:

- **Action** games have the highest median sales.
- **Shooter** and **Platform** genres have notable outliers indicating blockbuster hits.
- Most genres have a large range of sales, indicating both niche and mass-market games.

6. Regional Sales Insights

- 📍 **Regional Averages:**

Region	Avg Sales (in millions)
NA_Sales	~ 0.26
EU_Sales	~ 0.14
JP_Sales	~ 0.07
Other_Sales	~ 0.05

- North America leads in average sales, followed by Europe. Japan and other regions have significantly lower averages.

7. Top Publishers and Platforms

- **Publishers with most games**: Electronic Arts, Activision, Namco Bandai, Ubisoft
- **Platforms with most releases**: PS2, DS, PS3, Wii

8. Key Findings

- North America is the most lucrative market.
- Action and Sports dominate both in number and performance.
- Some genres (like Puzzle or Strategy) are niche but steady.
- Certain years (2006–2009) had a spike in releases — possibly due to console generation peaks.

9. Conclusion

This EDA provided insights into industry trends, helping to understand what genres and platforms dominated the market, where the biggest sales came from, and how trends evolved over time.
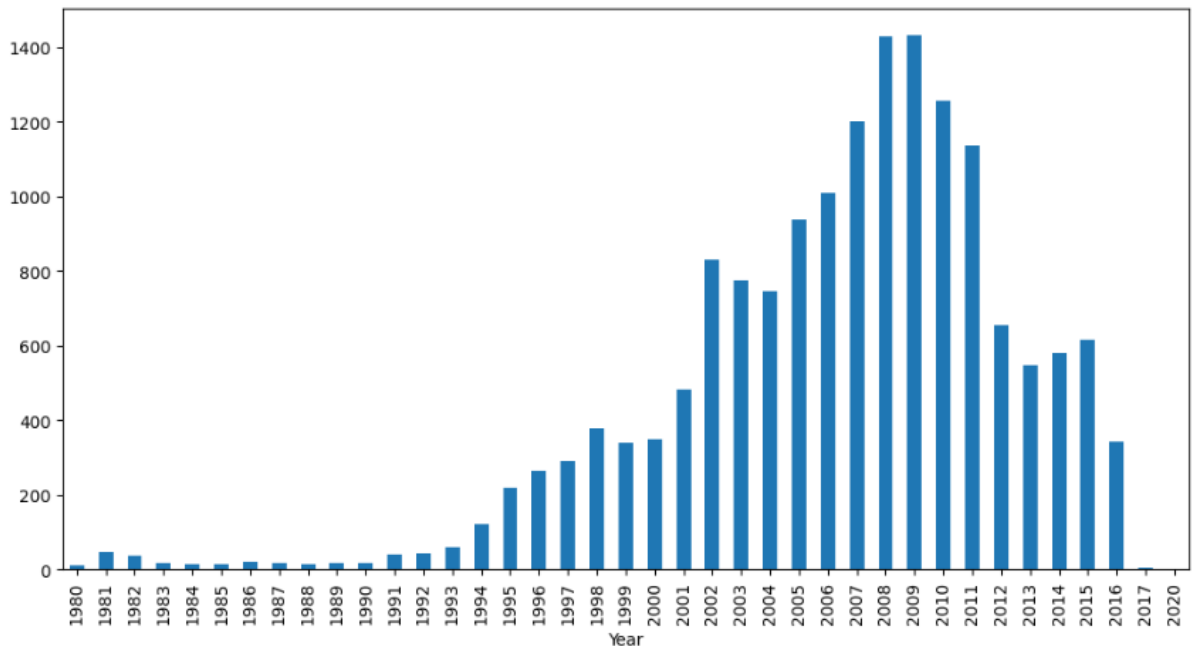

Fig: - Sales per year

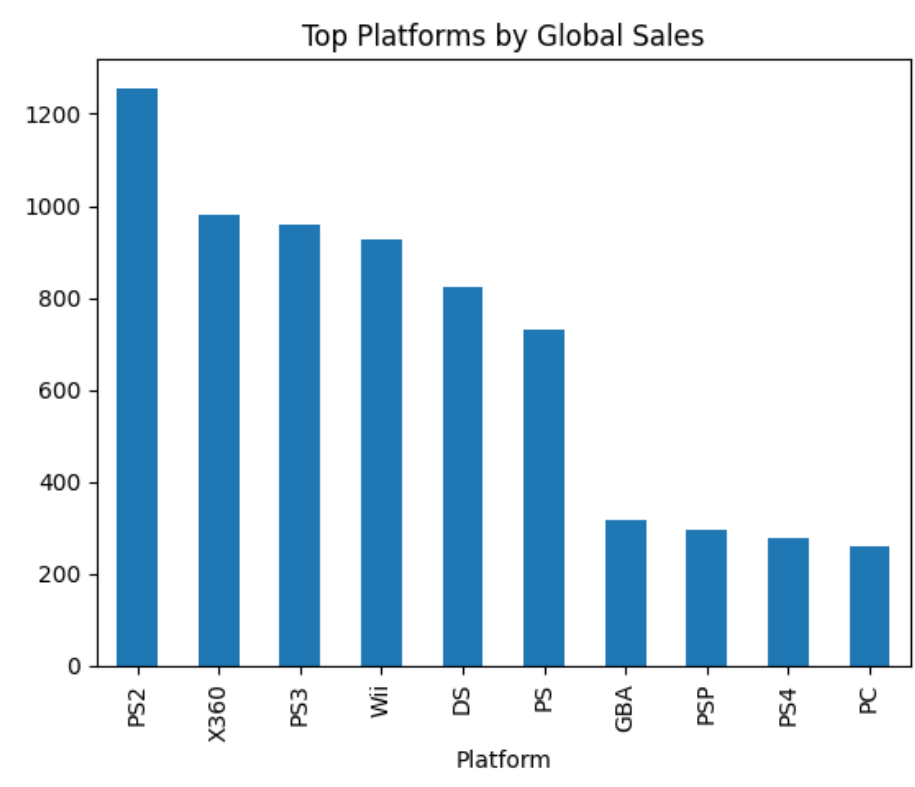|	Rank	Name	Platform	Year	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales
0	1	Wii Sports	Wii	2006.0	Sports	Nintendo	41.49	29.02	3.77	8.46	82.74
1	2	Super Mario Bros.	NES	1985.0	Platform	Nintendo	29.08	3.58	6.81	0.77	40.24
2	3	Mario Kart Wii	Wii	2008.0	Racing	Nintendo	15.85	12.88	3.79	3.31	35.82
3	4	Wii Sports Resort	Wii	2009.0	Sports	Nintendo	15.75	11.01	3.28	2.96	33.00
4	5	Pokemon Red/Pokemon Blue	GB	1996.0	Role-Playing	Nintendo	11.27	8.89	10.22	1.00	31.37
5	6	Tetris	GB	1989.0	Puzzle	Nintendo	23.20	2.26	4.22	0.58	30.26
6	7	New Super Mario Bros.	DS	2006.0	Platform	Nintendo	11.38	9.23	6.50	2.90	30.01
7	8	Wii Play	Wii	2006.0	Misc	Nintendo	14.03	9.20	2.93	2.85	29.02
8	9	New Super Mario Bros. Wii	Wii	2009.0	Platform	Nintendo	14.59	7.06	4.70	2.26	28.62
9	10	Duck Hunt	NES	1984.0	Shooter	Nintendo	26.93	0.63	0.28	0.47	28.31

Fig: - Top Selling Games Globally
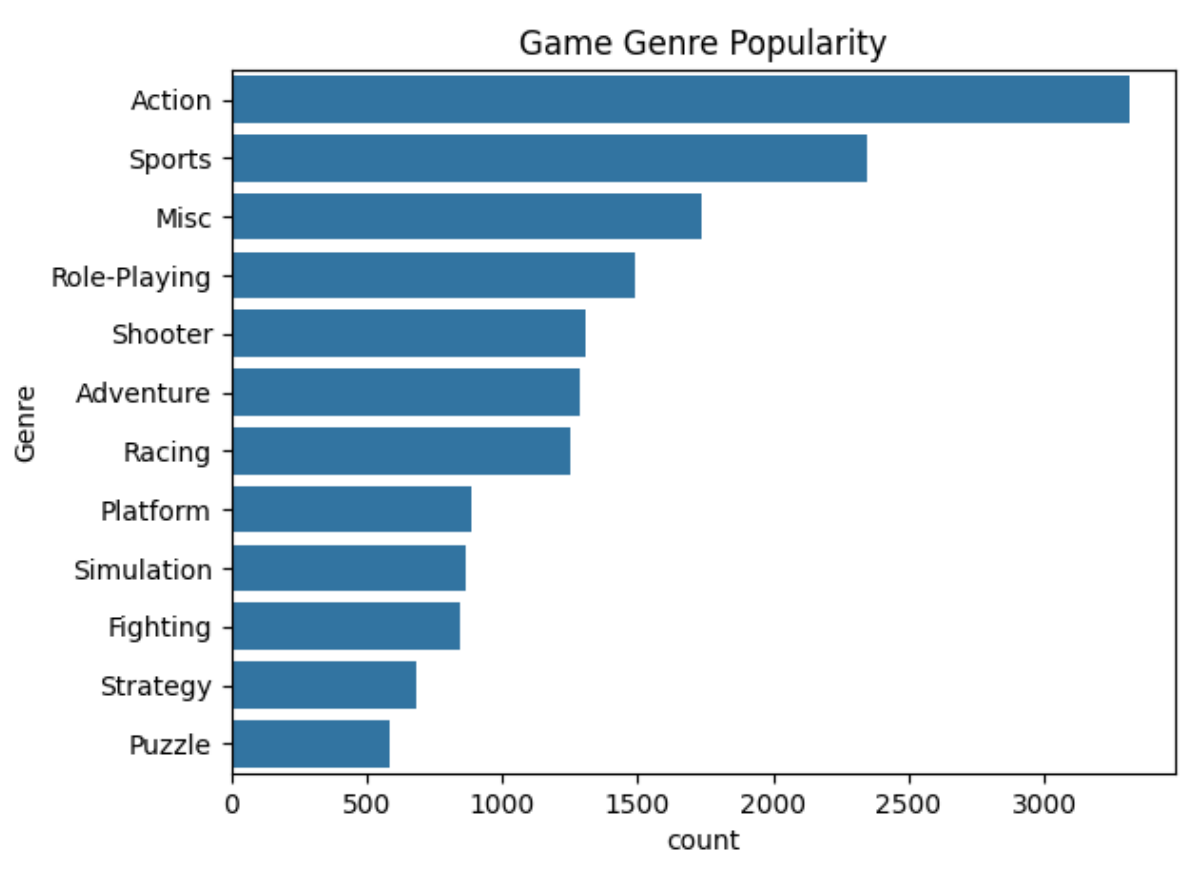
Fig: - Game popularity by Genre

Fig: - Top sales globally