

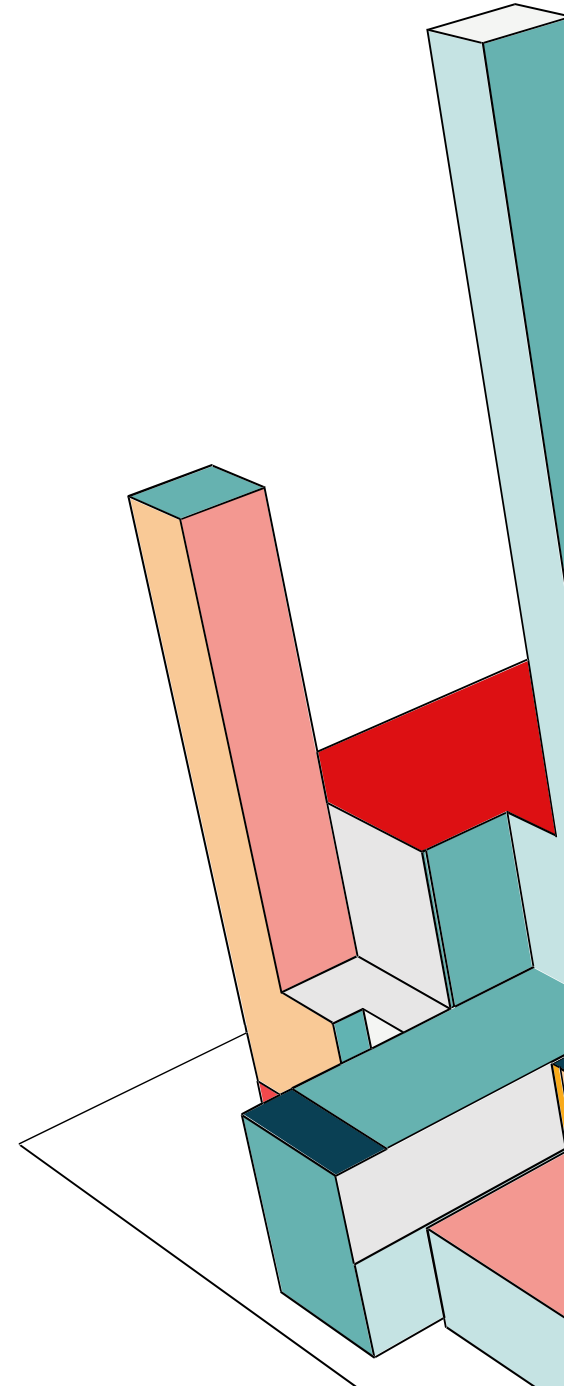


VIDEO GAME SALES PREDICTION PROJECT

Subash R

AGENDA

- Description of Dataset
- Project Goal
- Analysis and Findings
- Model Predictions
- Model Results
- Conclusion





DESCRIPTION OF DATASET

Video Game Sales Dataset

DESCRIPTION

The Video Game Sales data as of December 22, 2016 has total of 16 features and 16,720 samples data. There are global sales data available of video games in different platforms, different genre's in North America, Europe, Japan and other regions from 1980 to 2016.

The dataset encompasses various attributes such as game names, platforms, release years, genres, publishers, and more.

The goal is to uncover insights and trends within the video game industry, exploring factors that contribute to successful game sales and understanding market dynamics across different platforms and regions.



THE DATASET CONTAINS THE FOLLOWING FEATURES:

1. Name: Name of the video game
2. Platform: Platform of the video game (e.g., Wii, NES)
3. Year_of_Release: Year the video game was released
4. Genre: Genre of the video game (e.g., Sports, Platform)
5. Publisher: Publisher of the video game
6. NA_Sales: Sales in North America (in millions)
7. EU_Sales: Sales in Europe (in millions)
8. JP_Sales: Sales in Japan (in millions)
9. Other_Sales: Sales in other regions (in millions)
10. Global_Sales: Total global sales (in millions)
11. Critic_Score: Aggregate score compiled by Metacritic staff (out of 100)
12. Critic_Count: Number of critic reviews used in the Critic_Score
13. User_Score: Score given by users (out of 10)
14. User_Count: Number of user reviews
15. Developer: Developer of the video game
16. Rating: ESRB rating (e.g., E for Everyone, M for Mature)

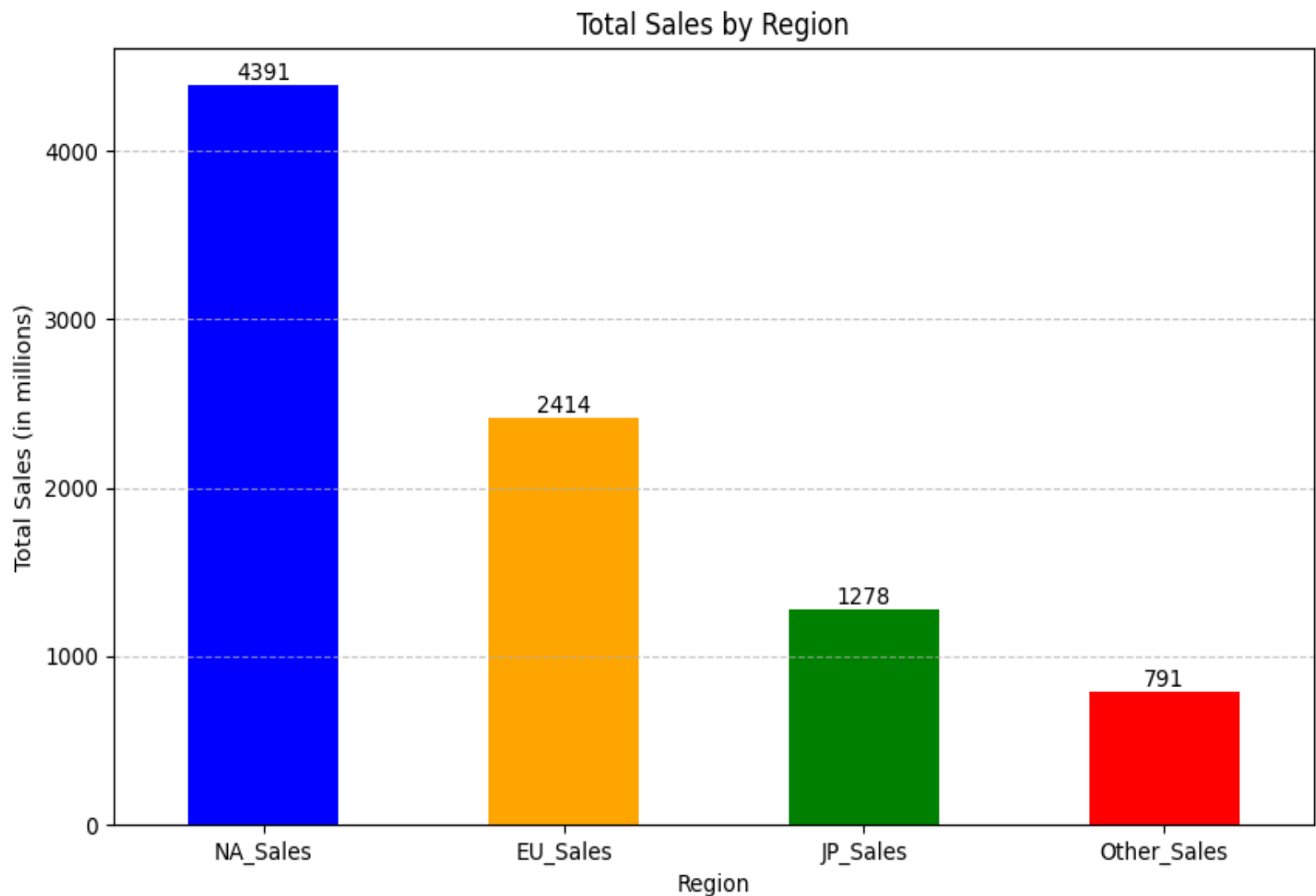




ANALYSIS & FINDINGS

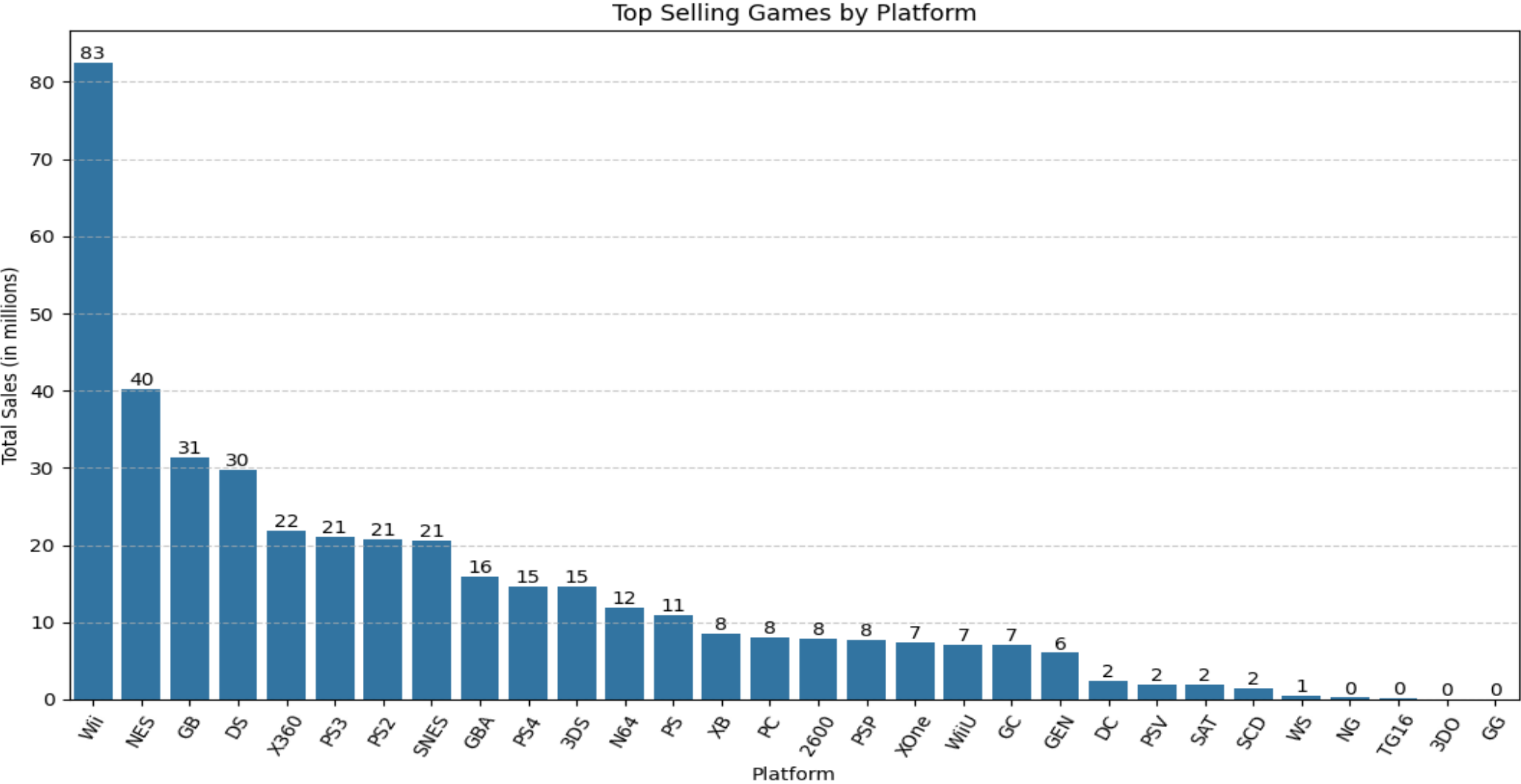
SALES ANALYSIS BY REGION

North America has the highest total sales, followed by Europe, Japan, and other regions.



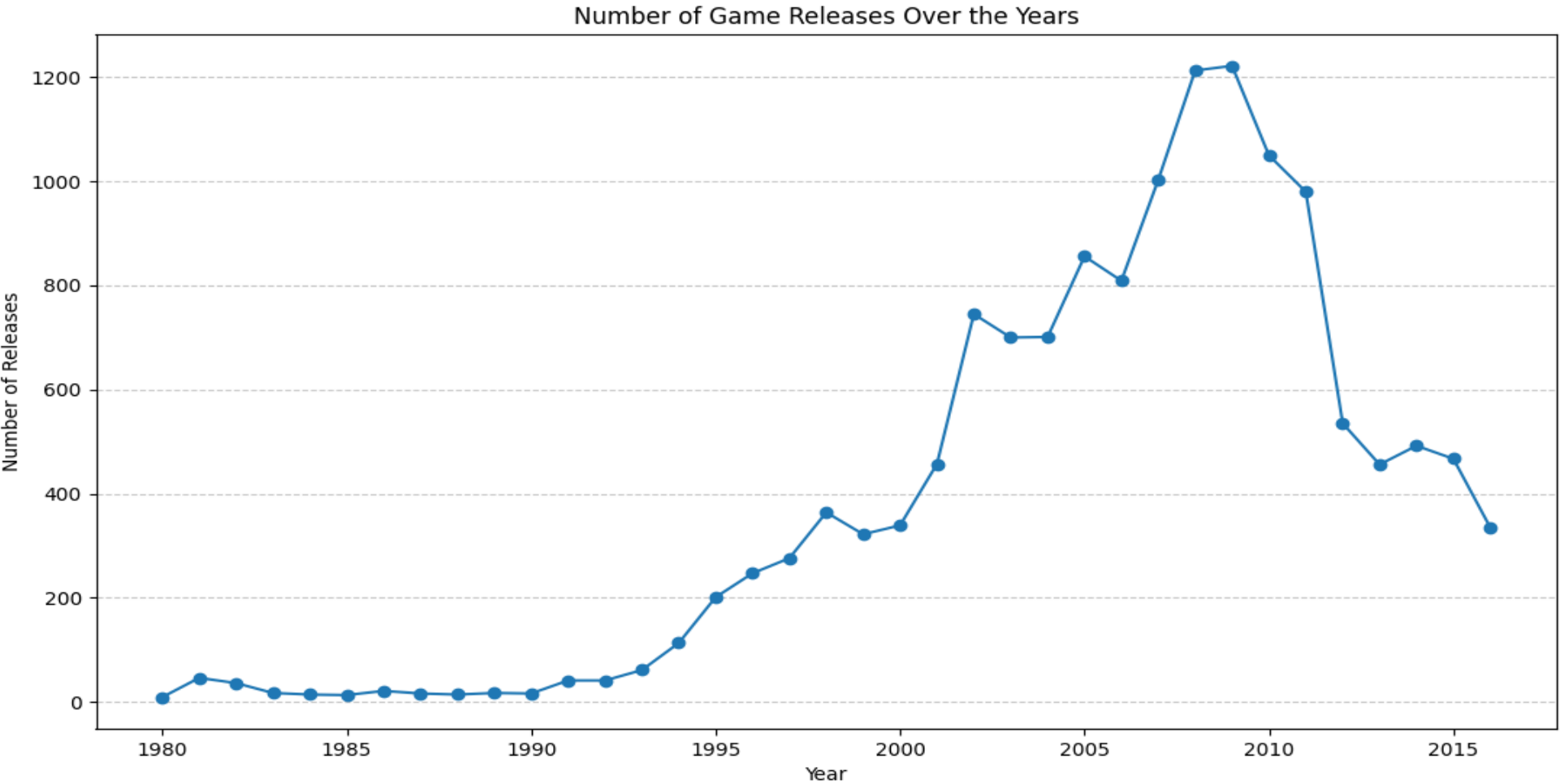
TOP-SELLING GAME BY PLATFORM

On Wii Platform has the highest sales of 83 millions over other platforms

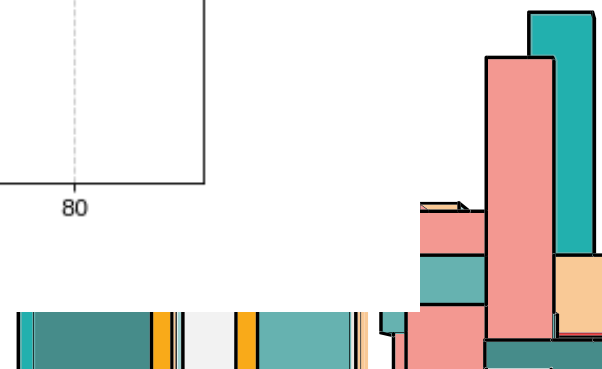
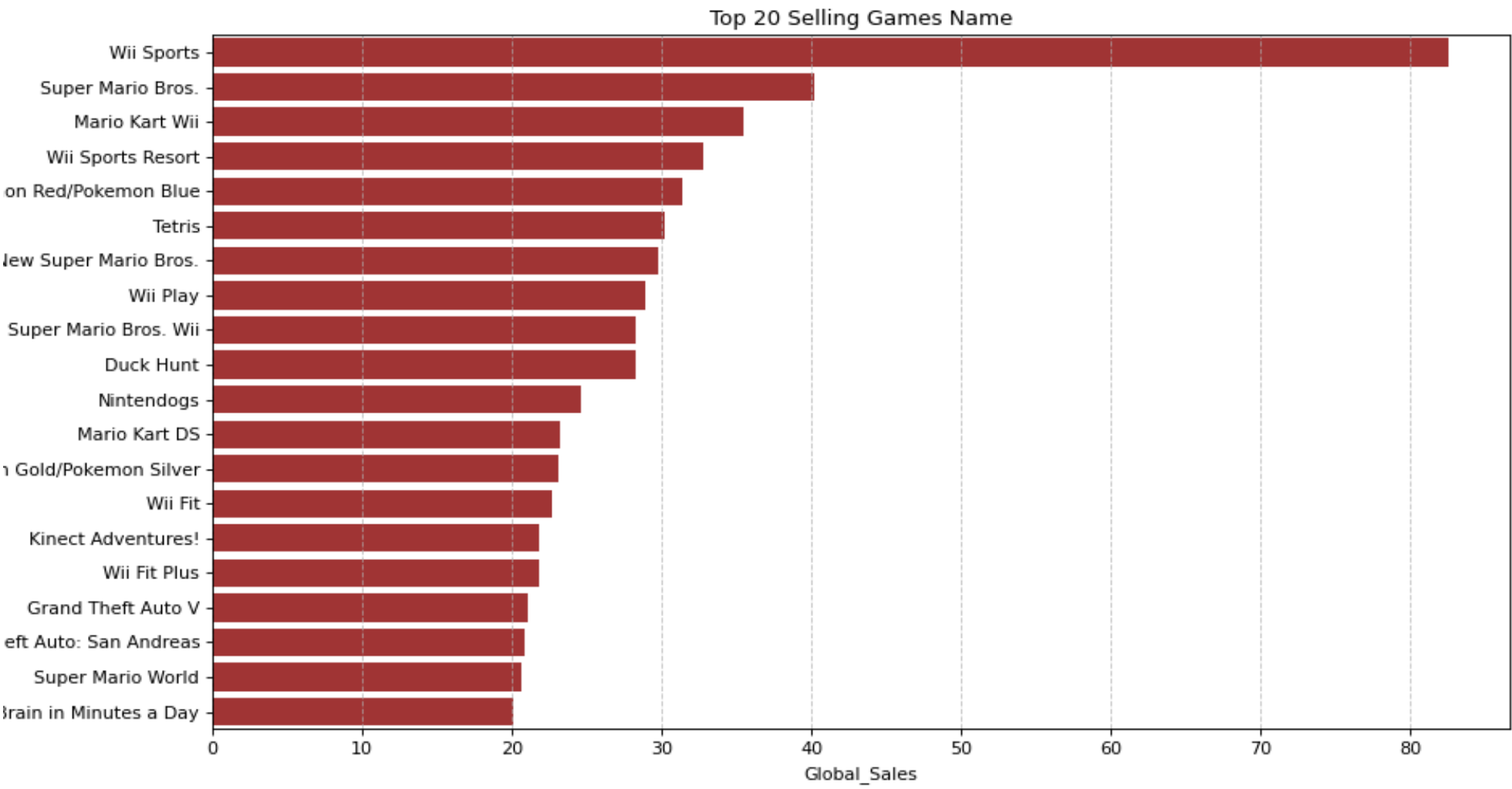


NUMBER OF GAME RELEASES OVER THE YEAR

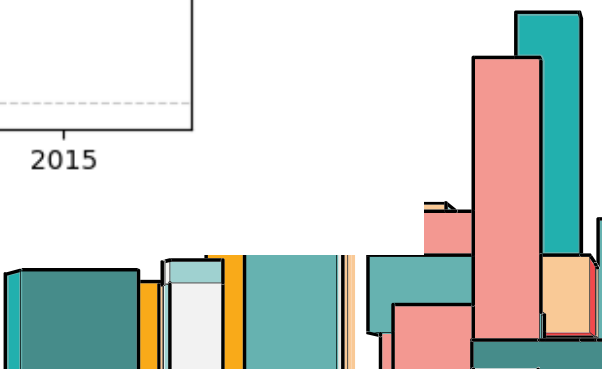
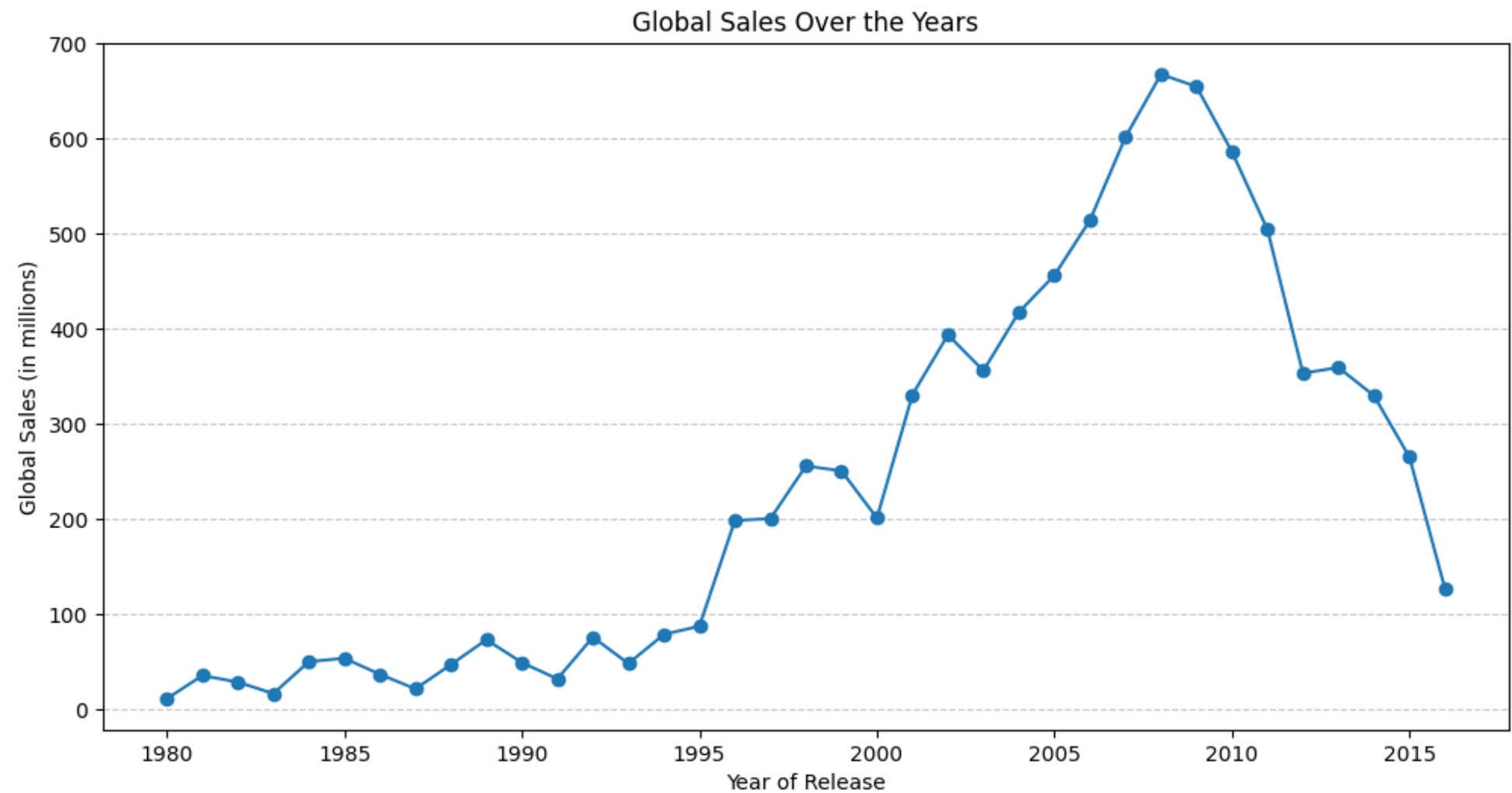
In 2005 to 2010 had the most game releases



TOP 20 SELLING GAME NAME'S

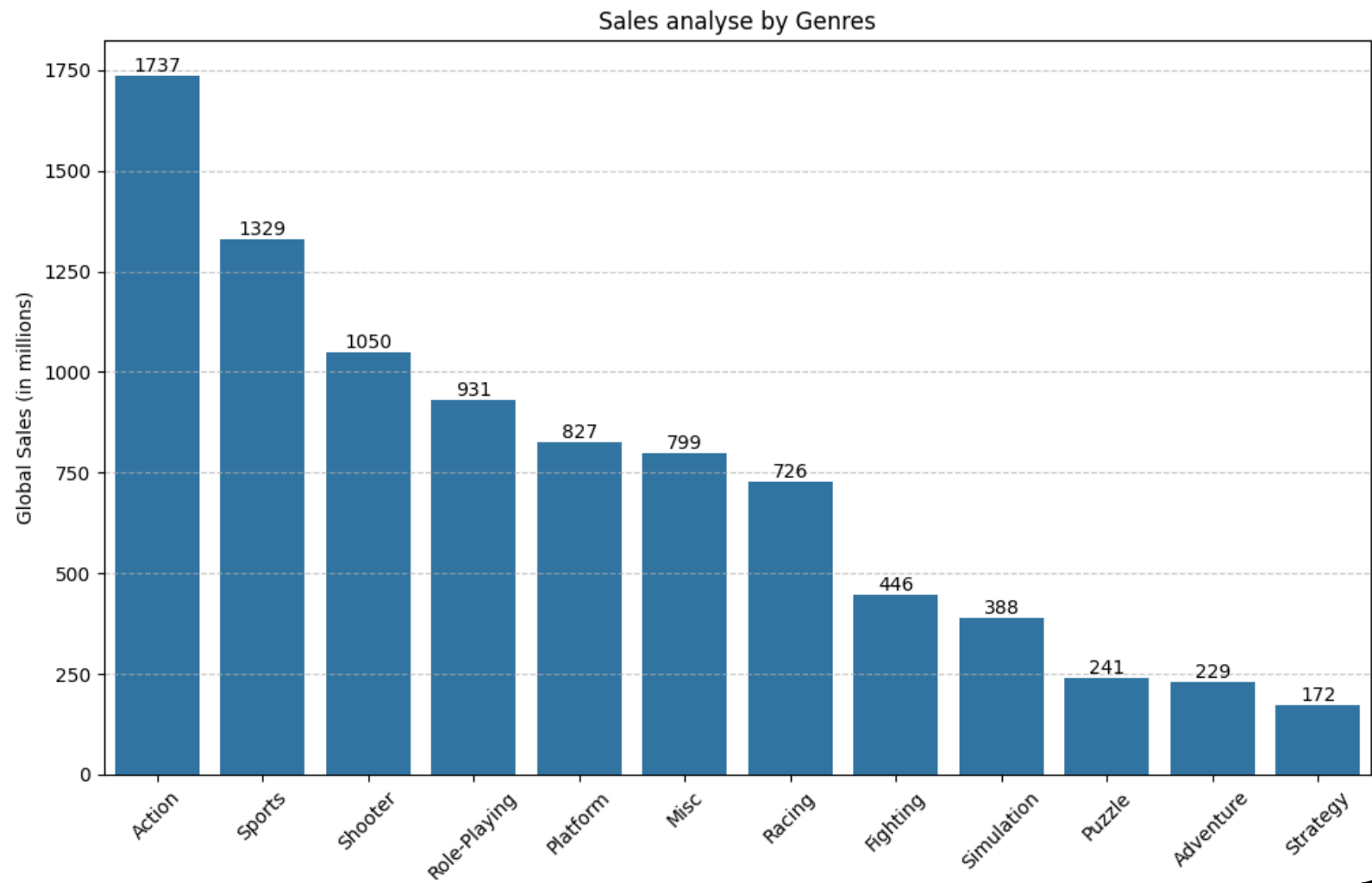


GLOBAL SALES OVER THE YEARS



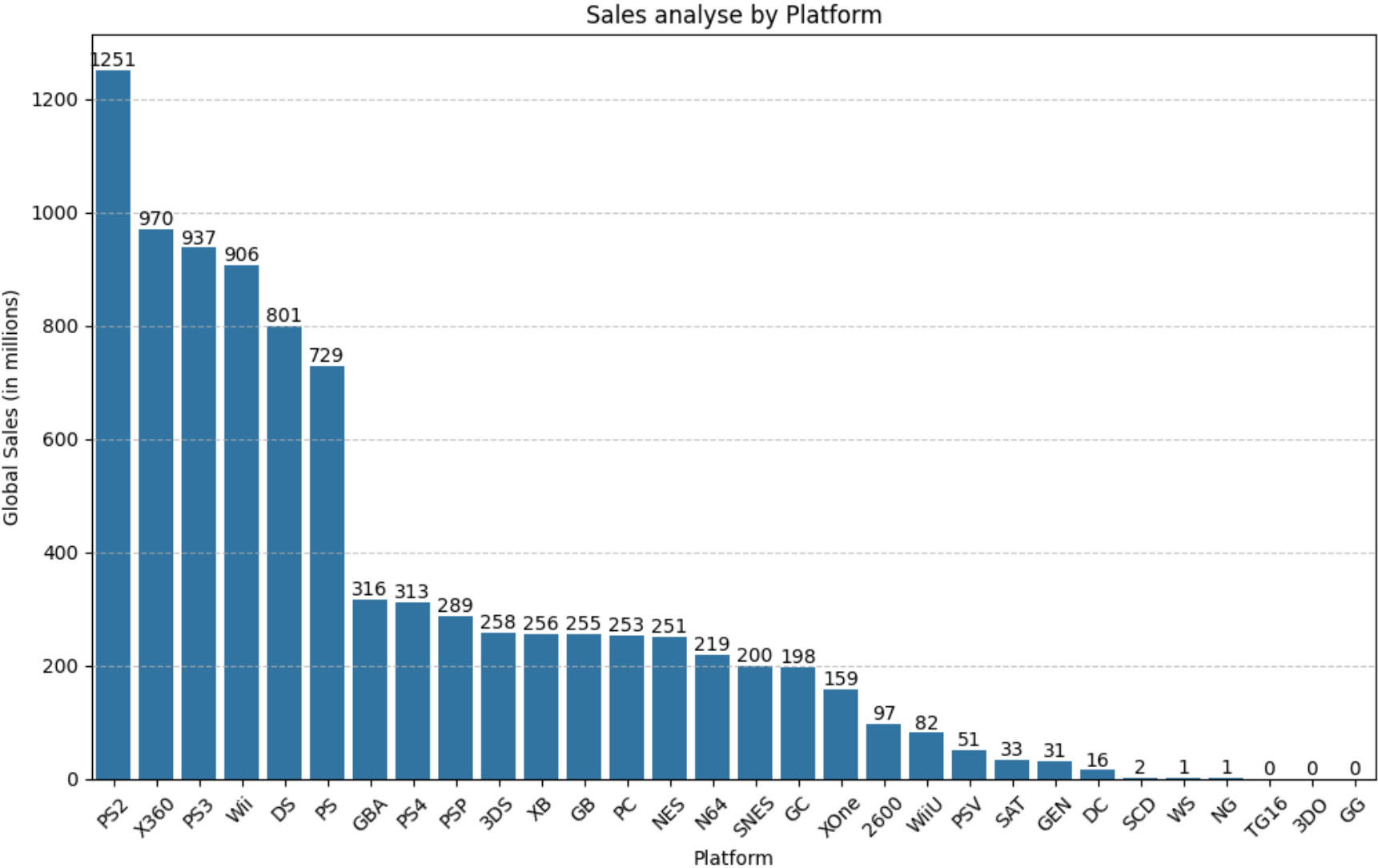
SALES ANALYSE BY GENRES

Action genre has the highest global sales of 1737 millions

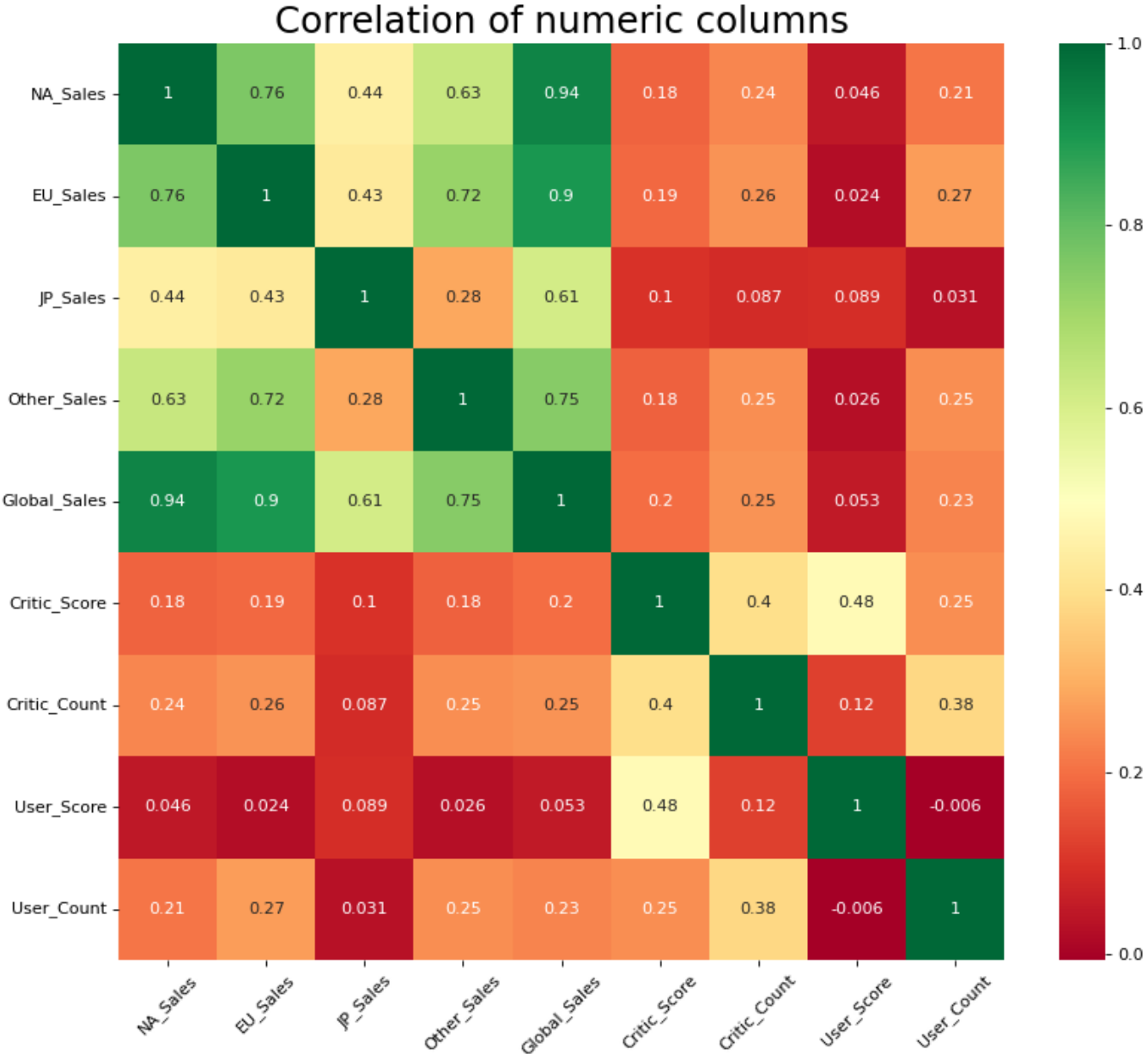


SALES ANALYSE BY PLATFORM

In PS2 platform has the highest global sales of 1251 millions



CORRELATION BETWEEN DATA

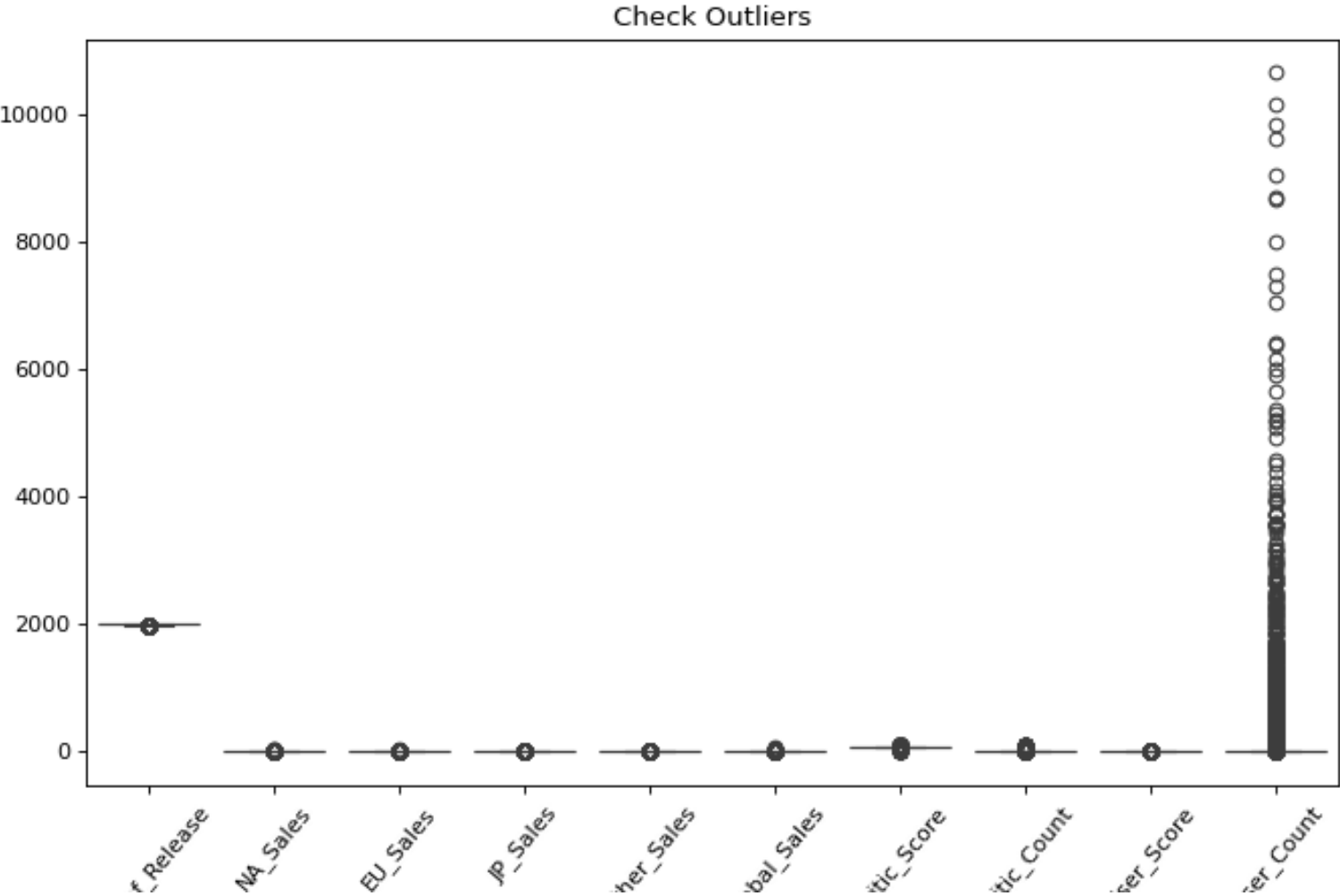


MODEL PREDICTIONS



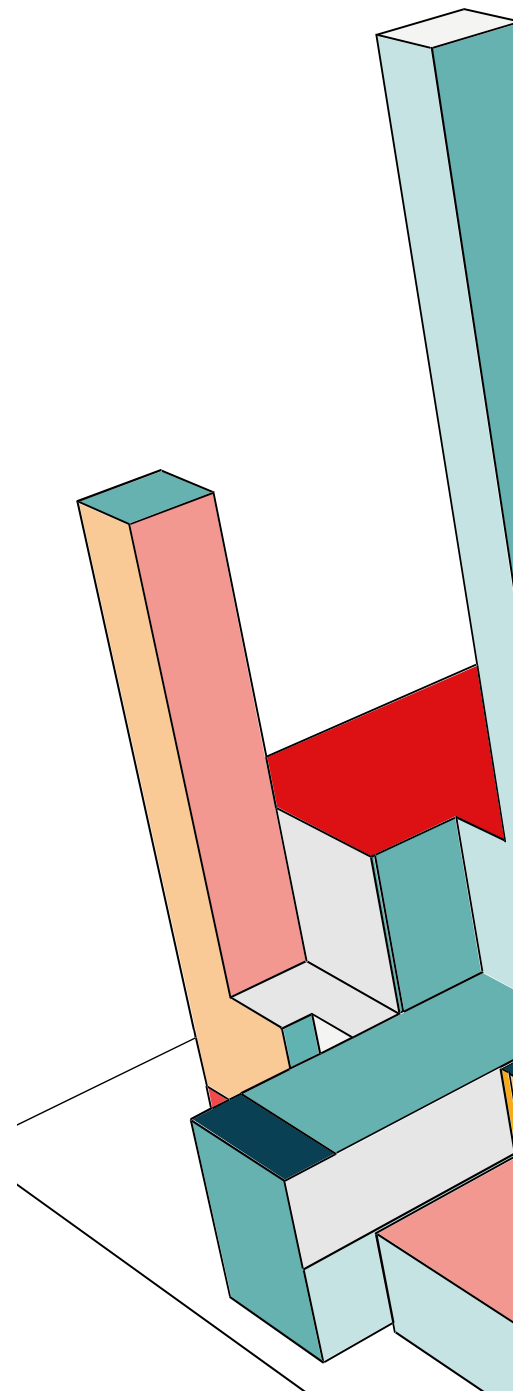
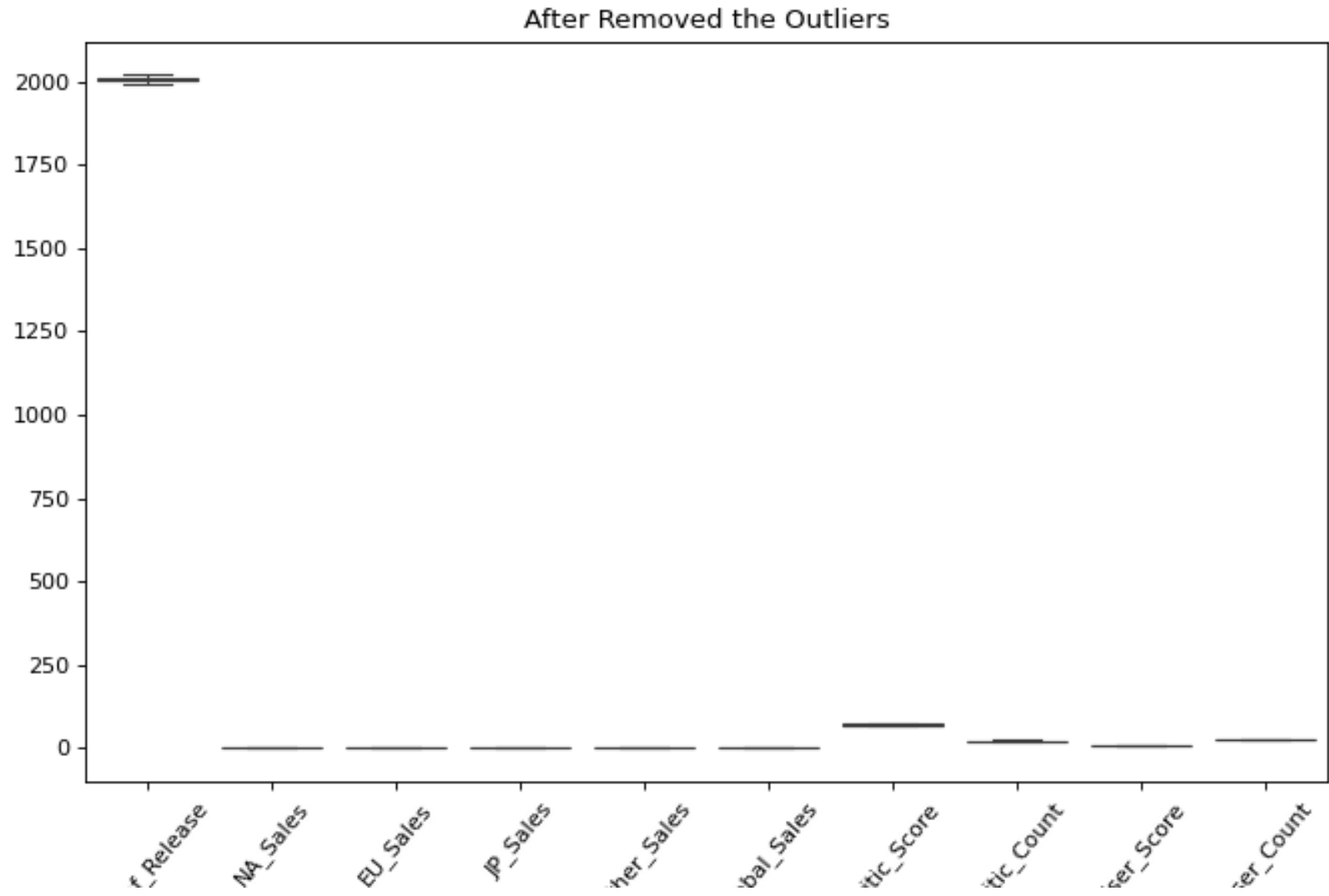
CHECK OUTLIERS FOR MODELS

Actual dataset (shape) : 14474 Samples, 16 features.

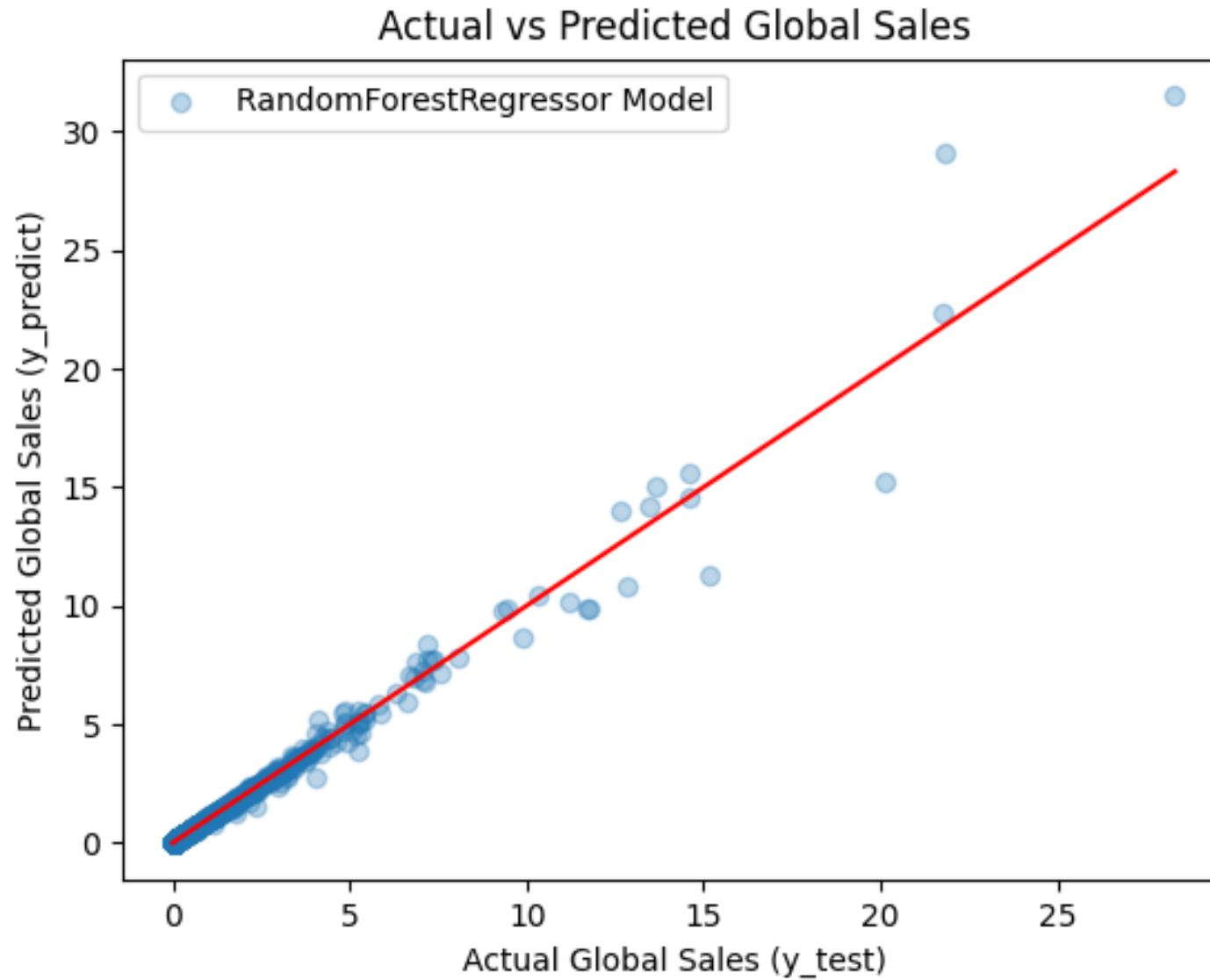


REMOVED OUTLIERS

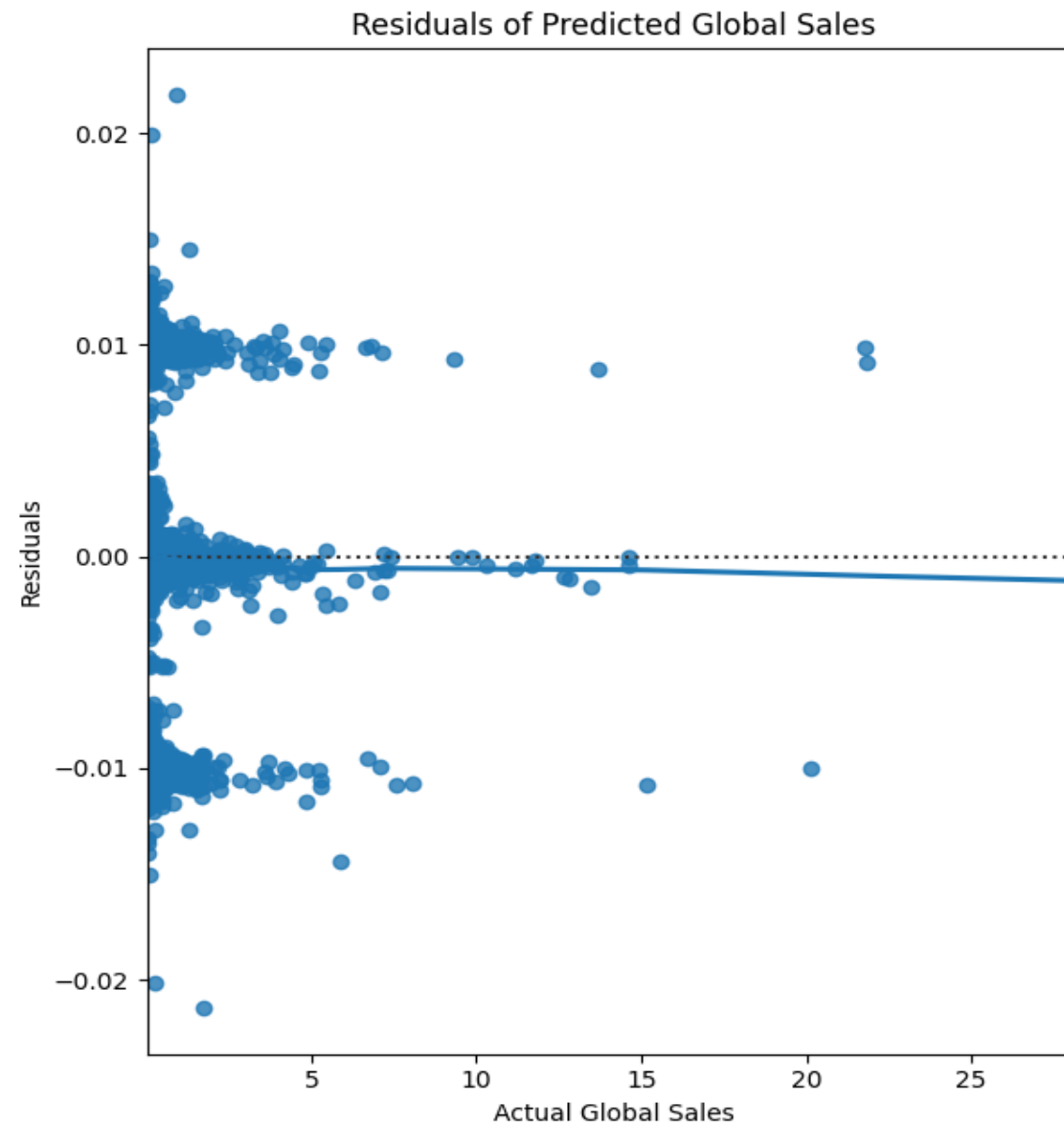
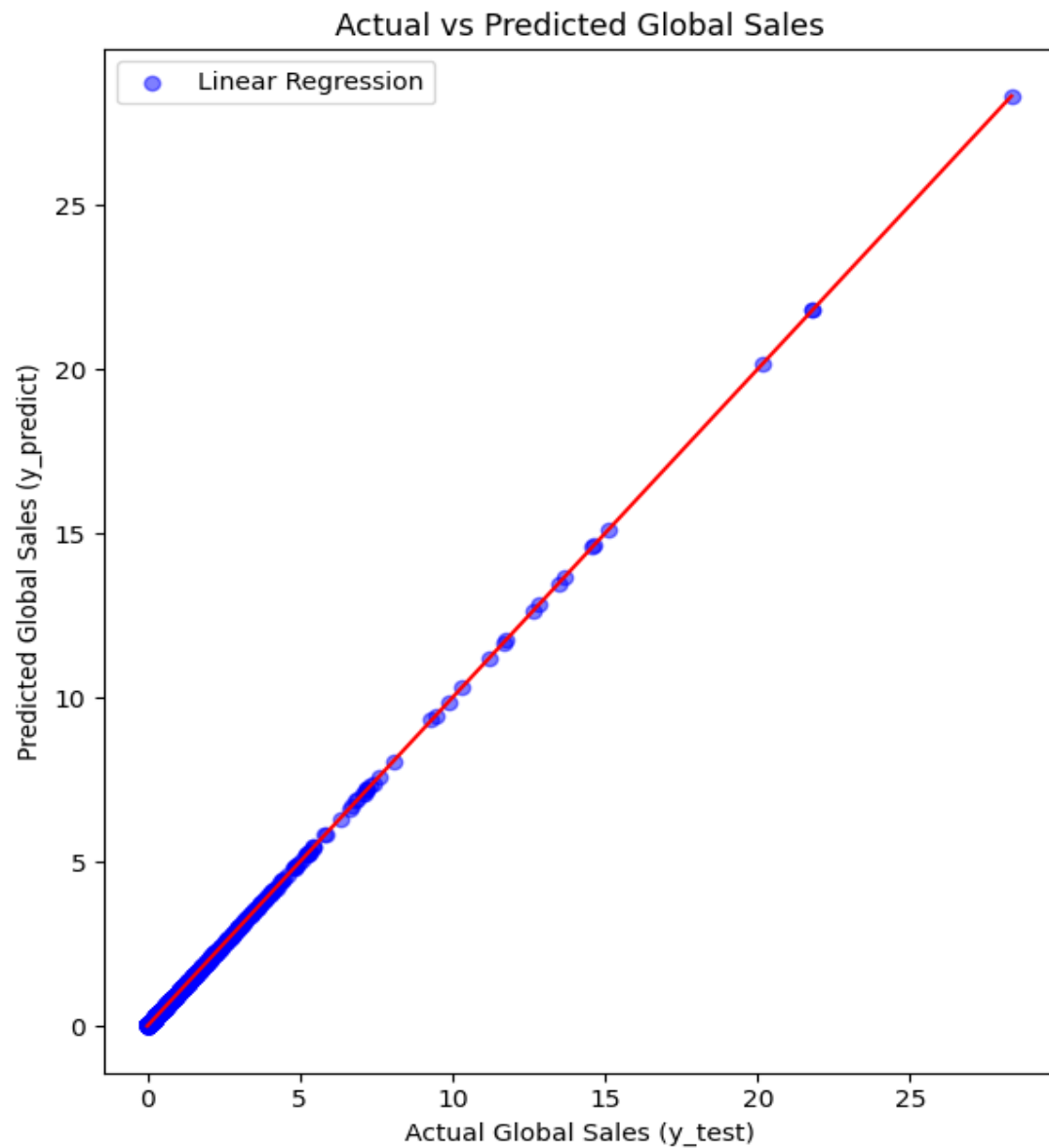
Removed outliers from the dataset : 10,199 Samples



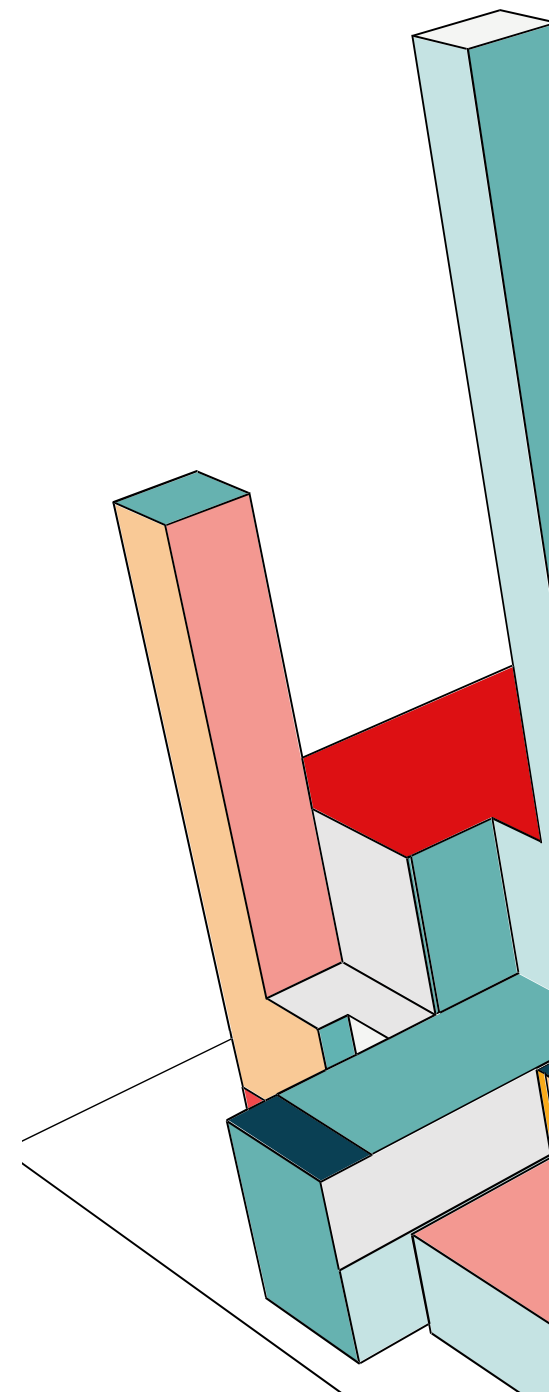
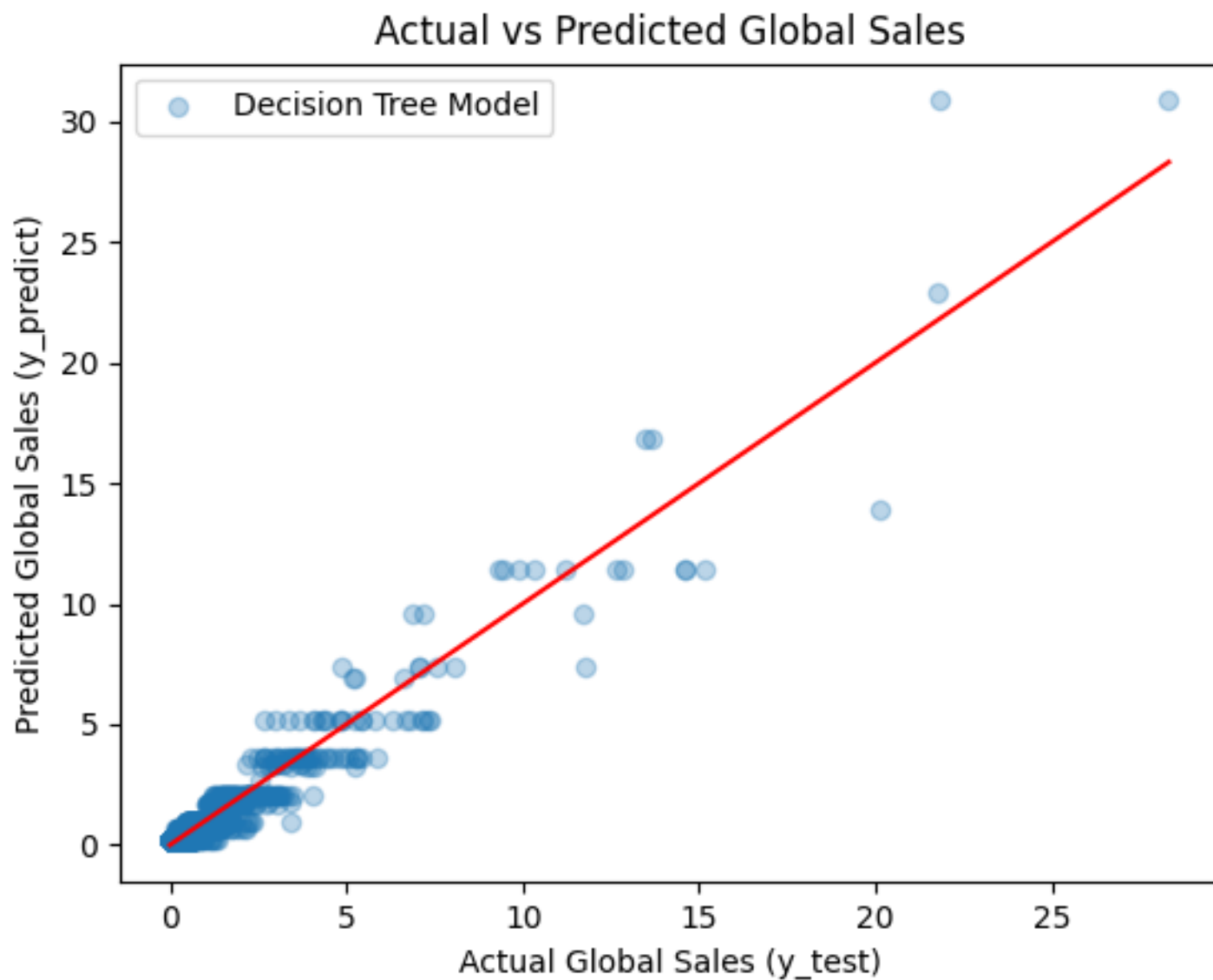
RANDOM FOREST REGRESSION



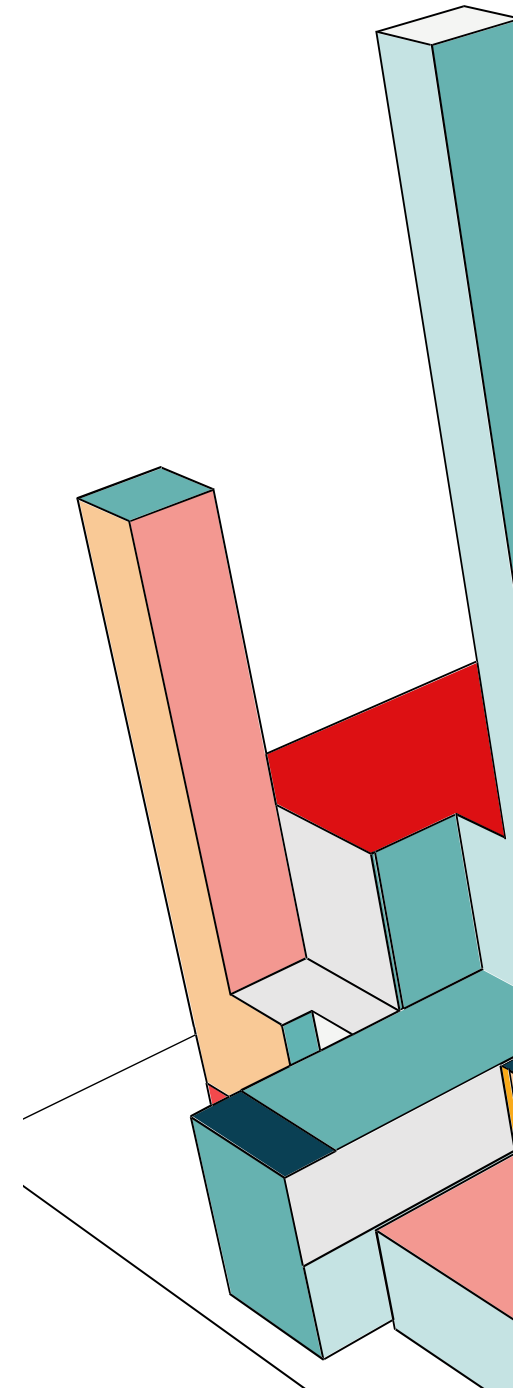
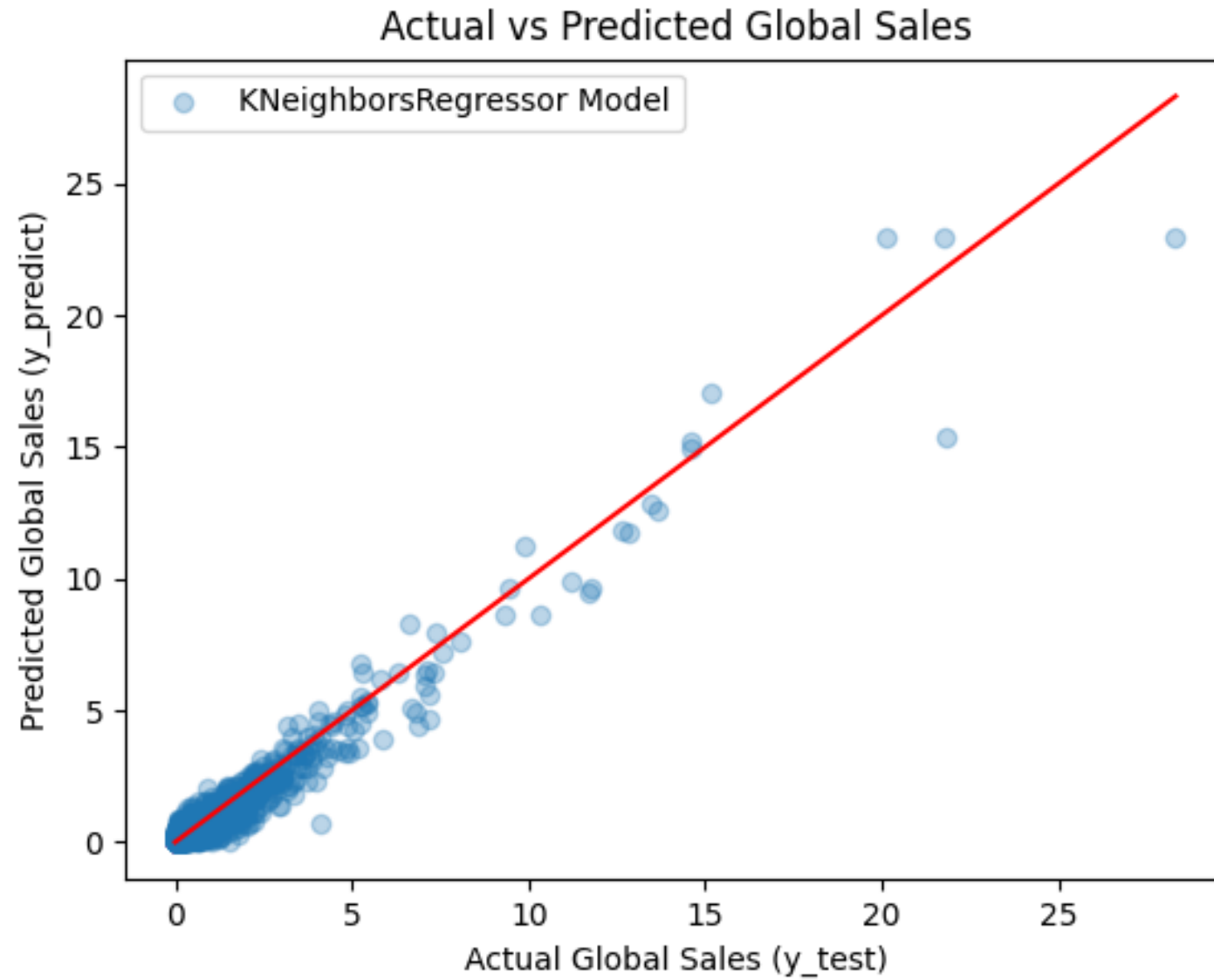
LINEAR REGRESSION



DECISION TREE REGRESSION



K-NEAREST NEIGHBOURS





MODEL RESULTS

R² SCORE OF DIFFERENT MODELS ARE:

R Squared Score value of Linear regression : 0.9999860217245073

R Squared Score value of Random Forest Regressor Model : 0.9779771009294983

R Squared Score value of Decision Tree Model : 0.9204874817308206

R Squared Score value of K-Neighbours Regressor Model : 0.9405905807674287



CONCLUSION

Each model has a good R^2 score, so they could be suitable for this project.

From the graph representation, we were able to learn about the findings in the data.

Examined the impact of different genres, platforms, and publishers on sales performance, and analysed regional differences in gaming preferences.

The findings highlight significant insights into the dataset.



THANK YOU

Subash R

