CSMODEL

PHASE ONE REPORT

SE

MP Group 4

Chu, Chuan-Chen
Jose, Yesha
Junsay, Andre
Mangawang, Felix

Dataset Description

Variable Descriptions of the dataset

Variable	Description
Name	Title of the video game
Platform	The gaming platform (e.g. Wii, NES, PS4)
Year_of_Release	Year when the game was released
Genre	Genre (category) of the game (e.g. Sports, Racing)
Publisher	Company that published the game
NA_Sales	Sales in North America (in millions of units)
EU_Sales	Sales in Europe (in millions of units)
JP_Sales	Sales in Japan (in millions of units)
Other_Sales	Sales in the rest of the world (in millions of units)
Global_Sales	Total worldwide sales (in millions of units)
Critic_Score	Average critic review score (typically 0–100)
Critic_Count	Number of critic reviews used to compute the critic score
User_Score	Average user review score (typically on a 0–10 scale)
User_Count	Number of user reviews used to compute the user score
Developer	Studio or company that developed the game
Rating	Age/content rating (e.g. ESRB ratings: E, T, M, etc.)

Video Game Sales and Ratings as of January 2017

- Dataset of 12,080+ Steam games with key details (price, reviews, playtime, etc.)
- Data from Steam API, SteamSpy, and custom scraper; only official games included
- Ownership and playtime are estimates and time-sensitive
- Contains missing/inconsistent data needing careful handling

Data Cleaning

Initial Setup with Pandas Python on Jupyter

The dataset pulled from Kaggle was put into a Jupyter notebook. Python, with libraries pandas, numpy, matplotlib, were used.

Platform grouping based on perspective company

All PlayStation consoles such as the PS, PS2, PS3, PS4 were grouped into the same PlayStation platform, same goes for the others.

Unknown Values for Ratings

Some games remained unrated and had empty cells in the data file. These unknowns were replaced to none.

EDA Questions

Exploratory Data Analysis

EDA focuses on sales success using Genre, Critic/User Scores, Platforms, and Publishers, analyzing their impact on game performance.

Q1 Game Genres

Which game genres are associated with higher global sales?

Q2 Critic-User Scores

How do critic and user scores relate to global sales?

Q3 Top Platforms

Which gaming platforms tend to have higher-selling games?

Q4 Top Publishers

Who are the top publishers in terms of global sales performance?

Game Genres

Shooter Genre Among the Best

Shooter games have the highest average global sales (excluding Miscellaneous), highlighting their broad appeal and dominant commercial success over other genres.

Average Global Sales by Genre
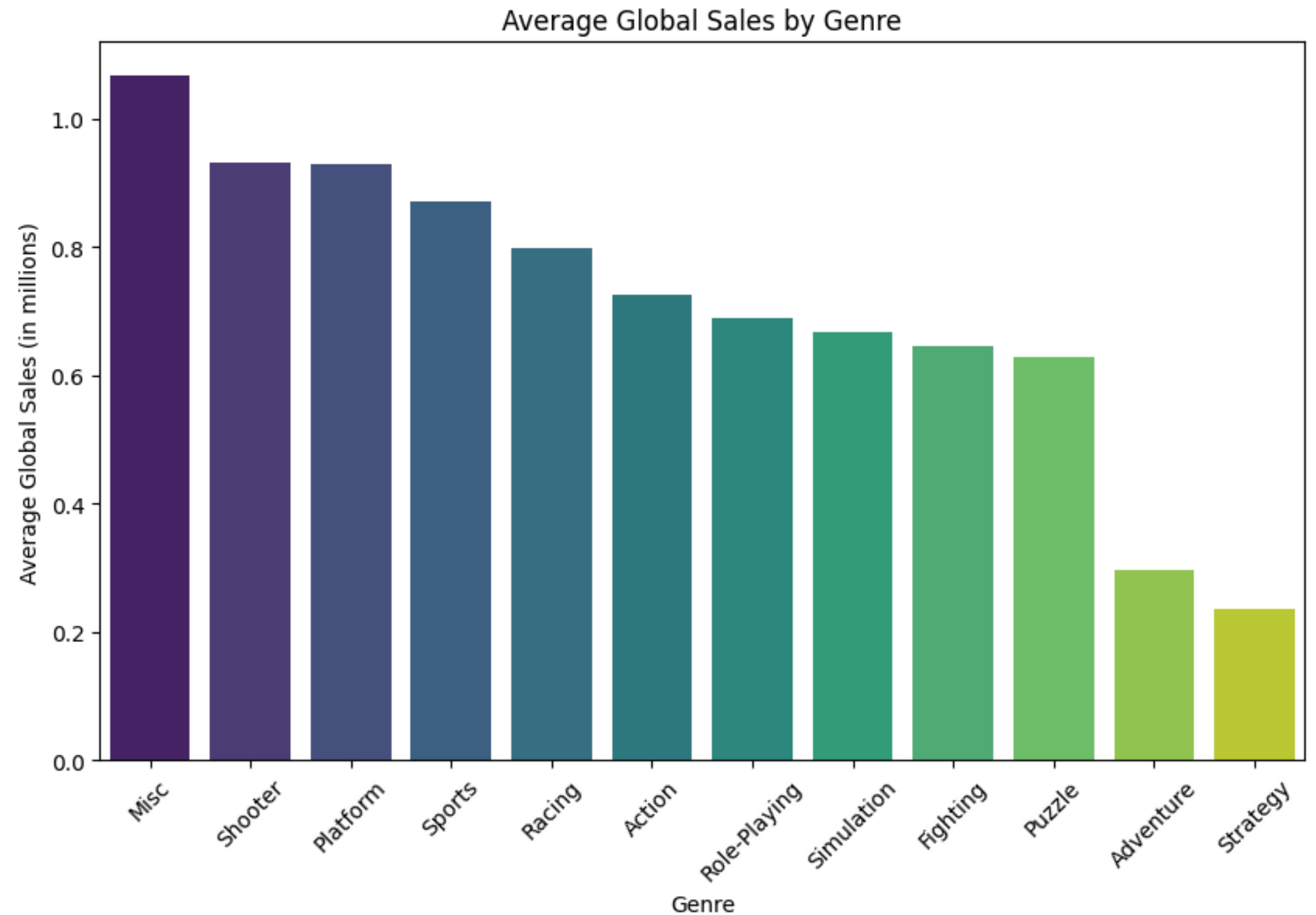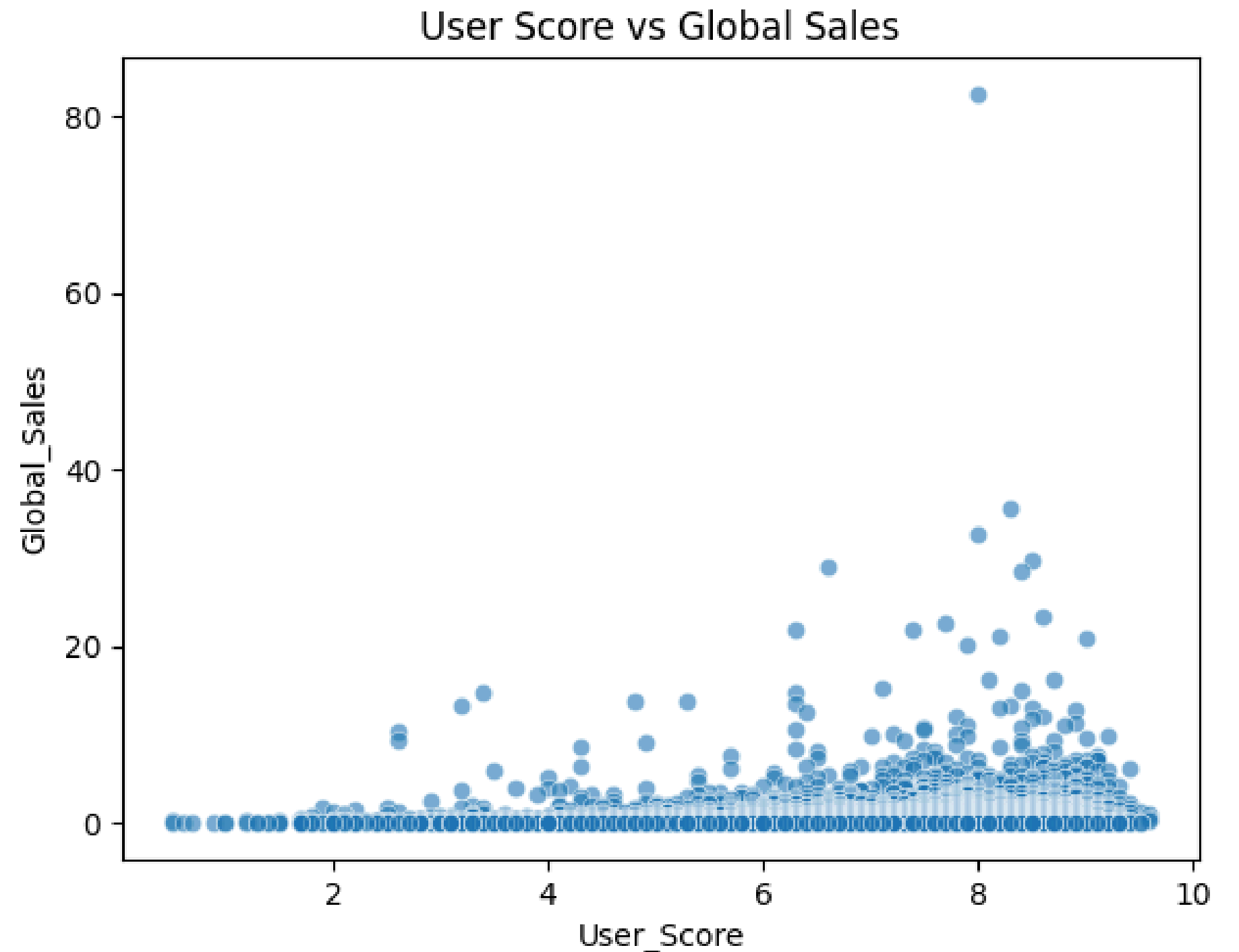
Critic-User Scores

Critic Score and Global Sales

There is a weak positive correlation between the two. Positive critic reviews are not the sole determinant for global sale performance.

Critic Score vs Global Sales
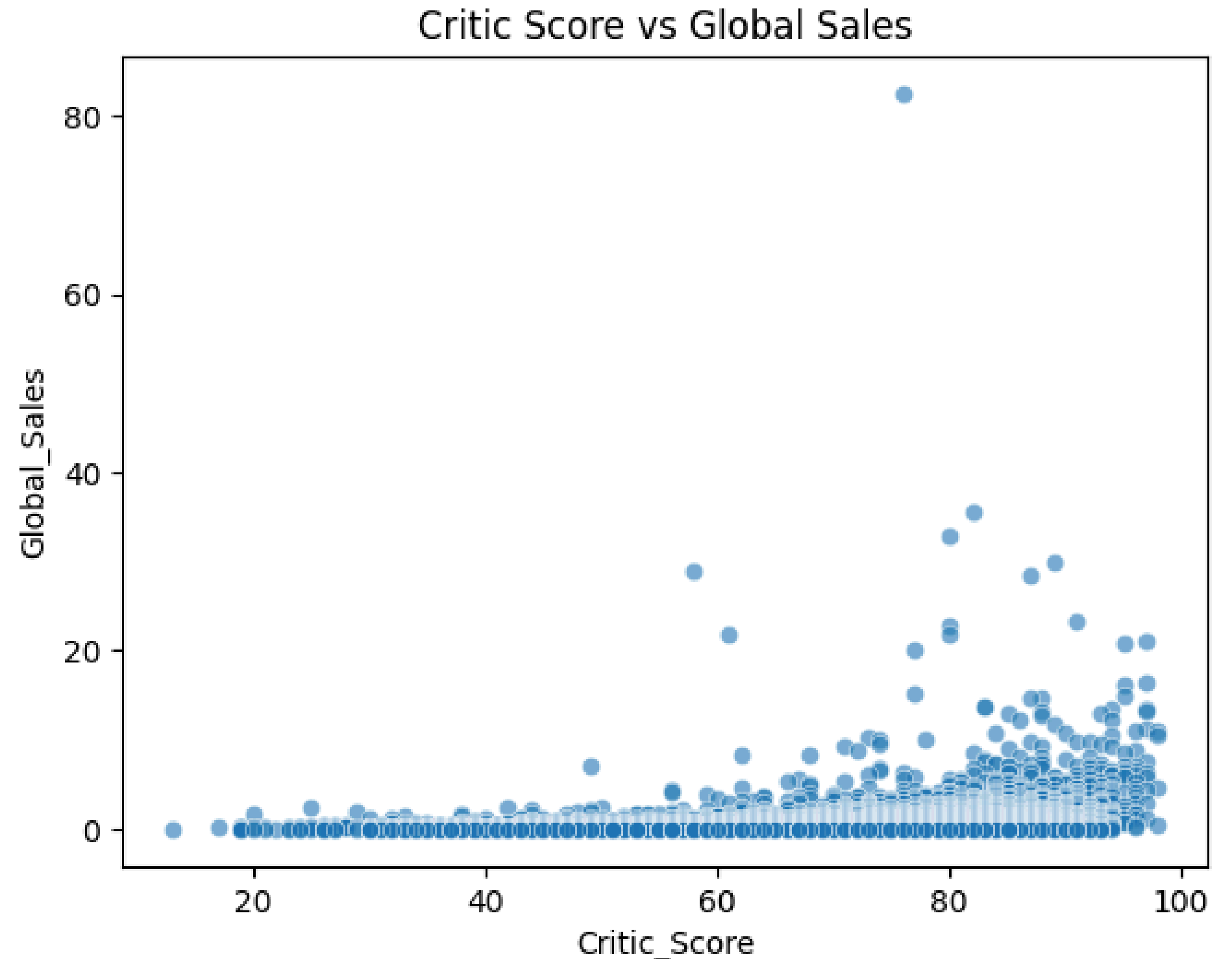
Critic-User Scores

User Score and Global Sales

The correlation between the two are even weaker. Player ratings have minimal direct influence on sales performance.

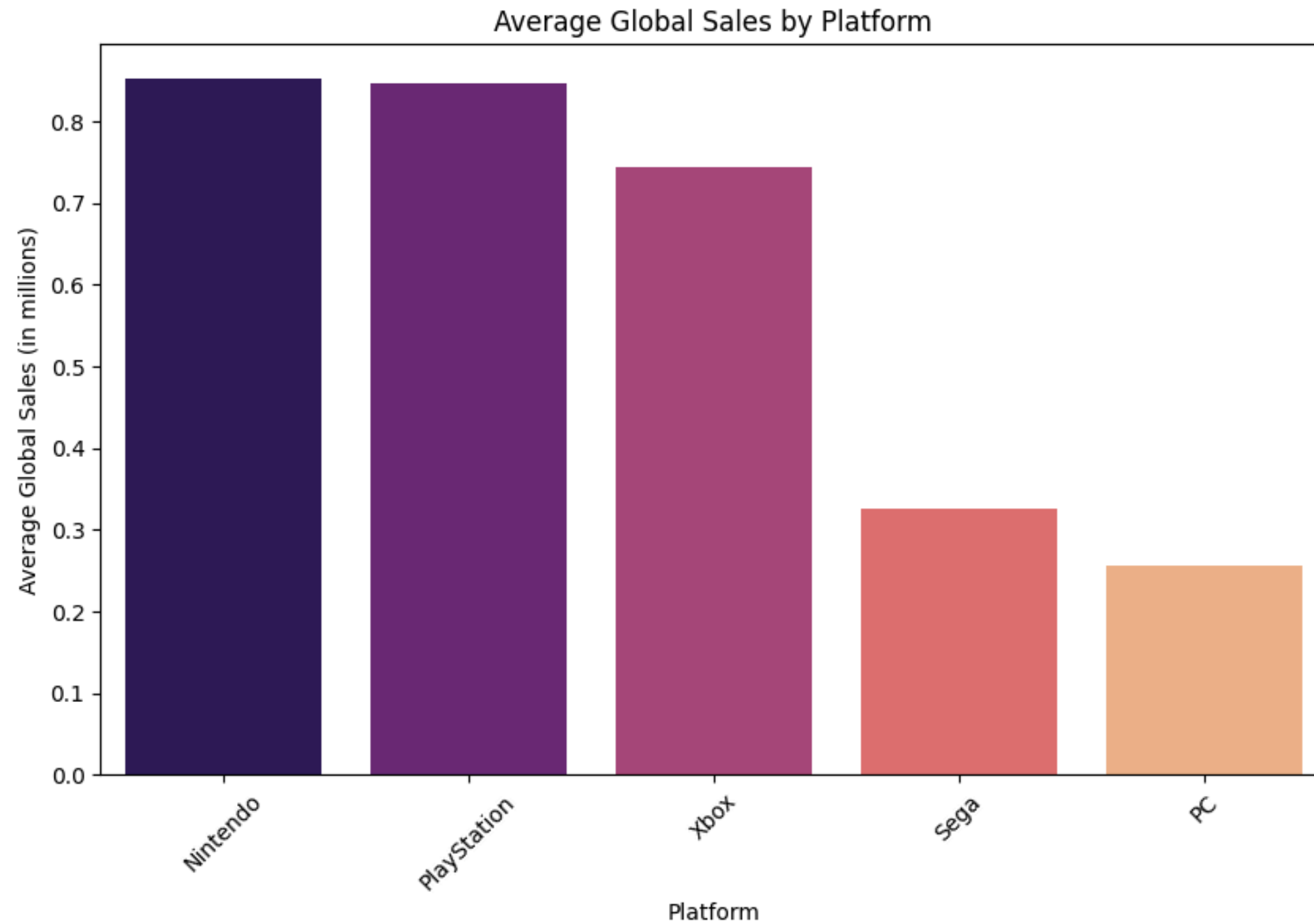

User Score vs Global Sales

Critic-User Scores

Critic-User Scores and Global Sales

Findings suggest well-reviewed games often achieve higher sales. Good reviews alone does not guarantee success.

Correlation Between Scores and Global Sales
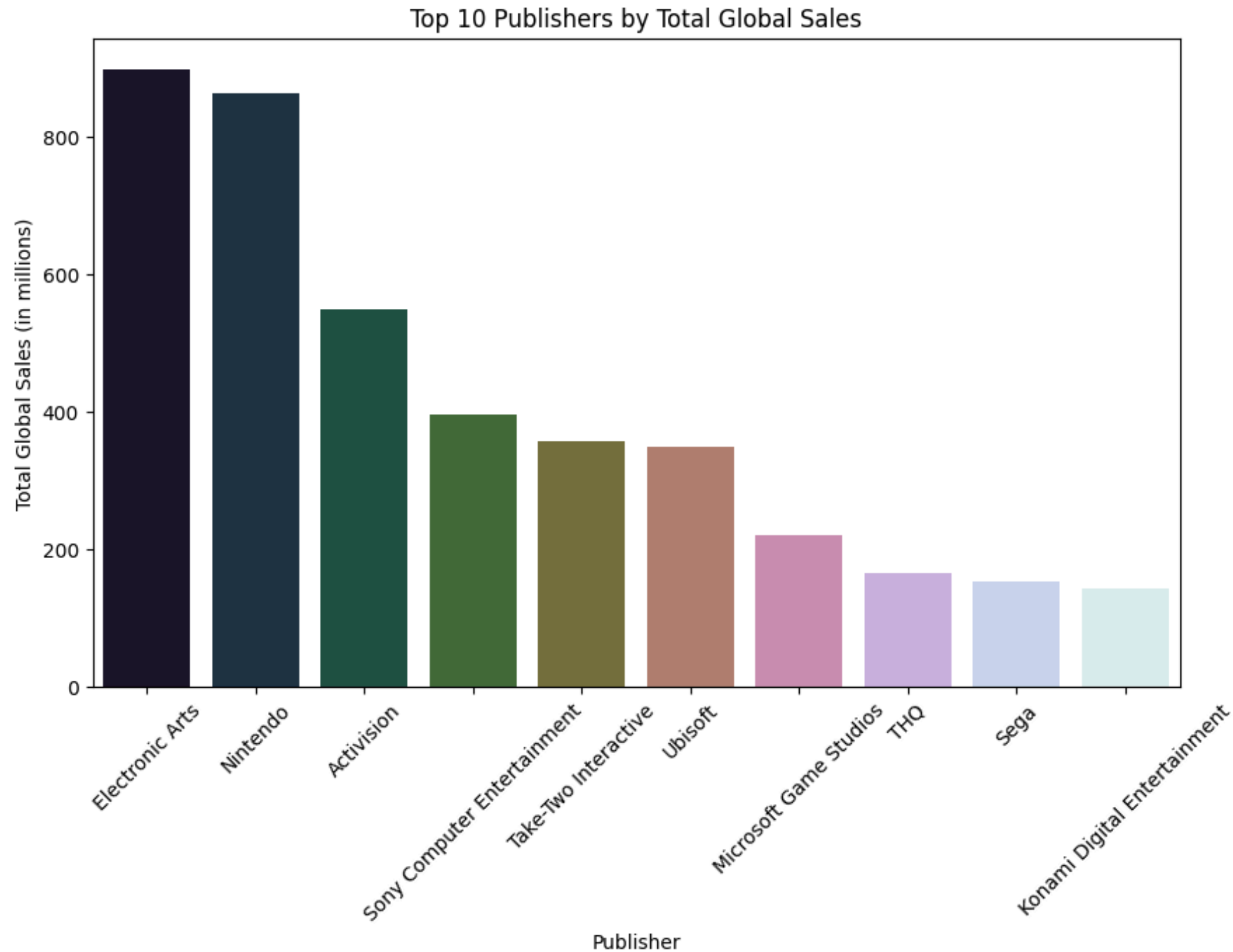
Average Global Sales by Platform
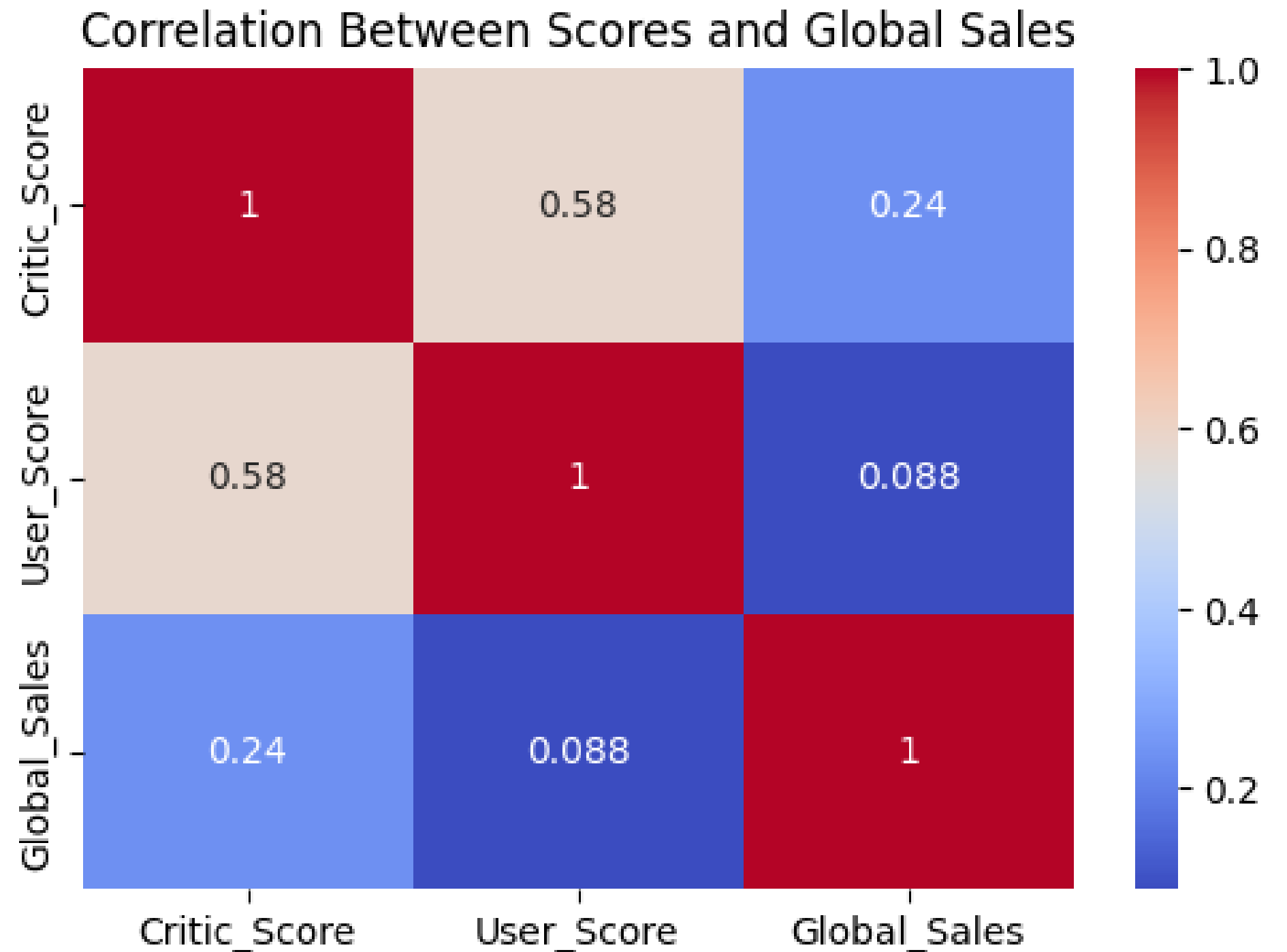
Best Platforms

Nintendo and PlayStation

Extremely close competition between the two with Nintendo holding only a slight edge in overall performance.

Top 10 Publishers by Total Global Sales

Top Publishers

Electronic Arts (EA) and Nintendo Leading

EA demonstrates strong dominance through franchises like FIFA, Battlefield, and the Sims. Nintendo follows suit with Mario, The Legend of Zelda, and Pokemon.

EDA Insights

- In Genre + Sales, tend to generate higher average and total sales.

- Scores + Sales, Moderate positive correlation with Critic Scores; weaker with User Scores.

- Platform + Sales, Platforms like PlayStation or Nintendo have higher-selling games on average.

- Publisher + Sales, EA, Nintendo and Activision account for a large share of total global sales.

Importance and Significance

What influences global game sales?

Key factors influencing global game sales include genre, critic/user scores, platform, publisher, and region.

How can game studios benefit from this?

This study explores their correlations to guide developers on effective marketing strategies.

What determines a game's success in terms of global sales?

How this study can benefit them?

An educated decision on the following example questions:

- Which publishing studio is the most competitive and yields the most global sales?

- How can we find the most success if the game studio aims for a region-based localization approach.
- (Ex. Japanese game studio aiming to release games only in Japan.)

- What platforms do this game's genre see more success in?

What determines a game's success in terms of global sales?

THANK YOU

MP Group 4

Chu, Chuan-Chen
Jose, Yesha
Junsay, Andre
Mangawang, Felix