Ria Vellanki

Trends in the Video Game Publishers

Introduction

The contemporary video game market is expansive, encompassing a diverse range of genres and platforms. As the number of games continues to grow, so too does the number of video game publishers, ranging from innovative indie developers to established industry giants. In this project, I explored the factors that contribute to a video game publisher's success. Key questions guiding this analysis included, which publishers produce the highest number of games? Is there a correlation between the number of games a publisher releases and their financial success? How does the genre of a game impact its sales? I found my dataset on Kaggle, the data focuses on video game sales and includes features such as Rank, Name, Publisher, Genre, Year, Global Sales, EU Sales, NA Sales, JP Sales, and Other Sales. You can find the dataset at this link: https://www.kaggle.com/datasets/gregorut/videogamesales. In the project, I focused on the Name, Publisher, Genre, Year, and Global Sales features in order to determine what makes a video game publisher successful on a large scale.

Data Preprocessing

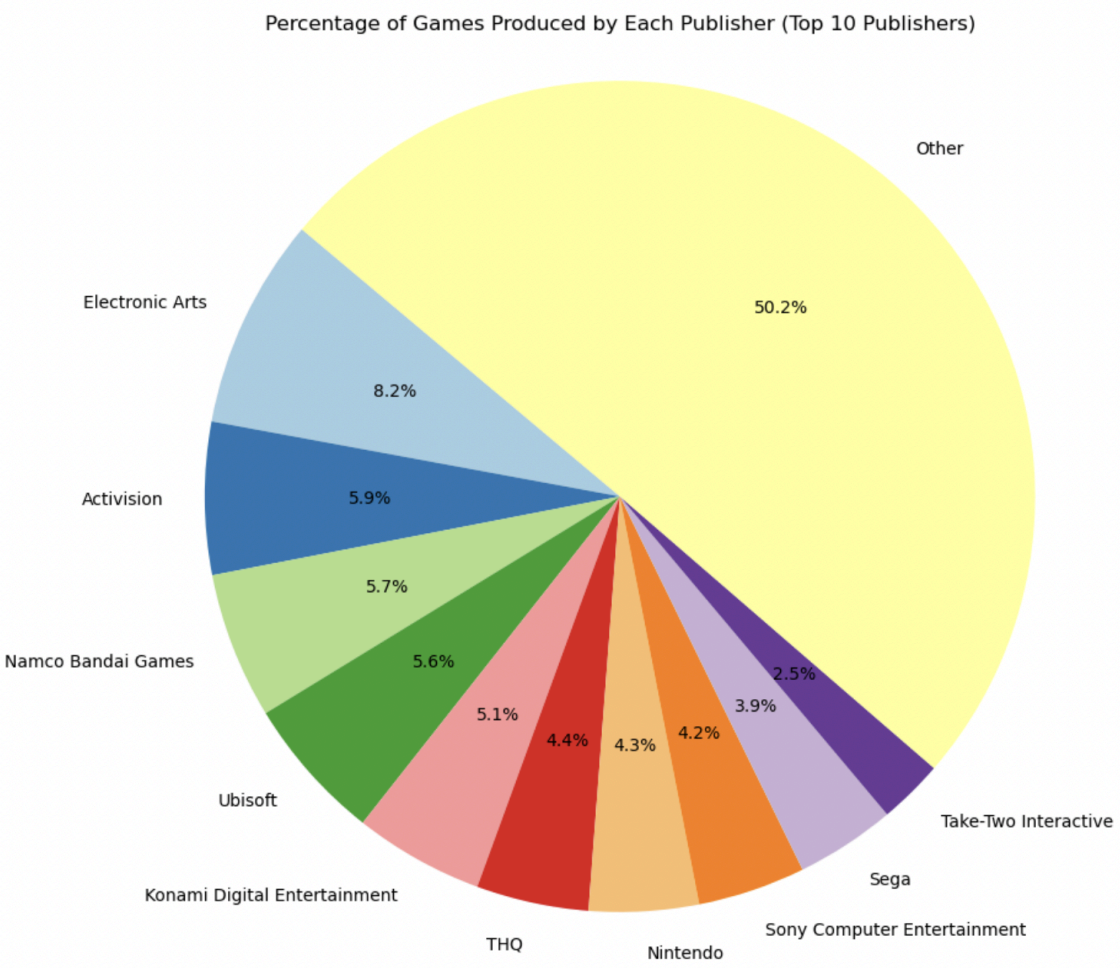
After I chose my dataset, I went into the data preprocessing stage of the data mining pipeline. I focused on two main methods of data preprocessing, removing null values and removing duplicates. I checked for null values using print(df.isnull().sum()), this line of code showed me the number of null values in each column. None of the columns had null values except for year and publisher. Year had 271 null values and

publisher had 58 null values. Using df.dropna(inplace=True), I removed all rows that contained a null value in order to maintain data integrity. Next, I used print(df.duplicated().sum()) to check for duplicate entries. There were 0 duplicates, therefore there was no need to go further with this method. I also thought about removing some columns from the dataframe as I planned to focus on global sales rather than regional sales but ultimately decided not to in the case that I decide to steer the focus of the project in a different direction later on.

Visualizations and Storytelling

My first visualization is a pie chart that shows how many games are made by each publisher. Since this data set has thousands of entries with hundreds of different publishers, I used the top 10 publishers with the most games and visualized the rest under the label of other.

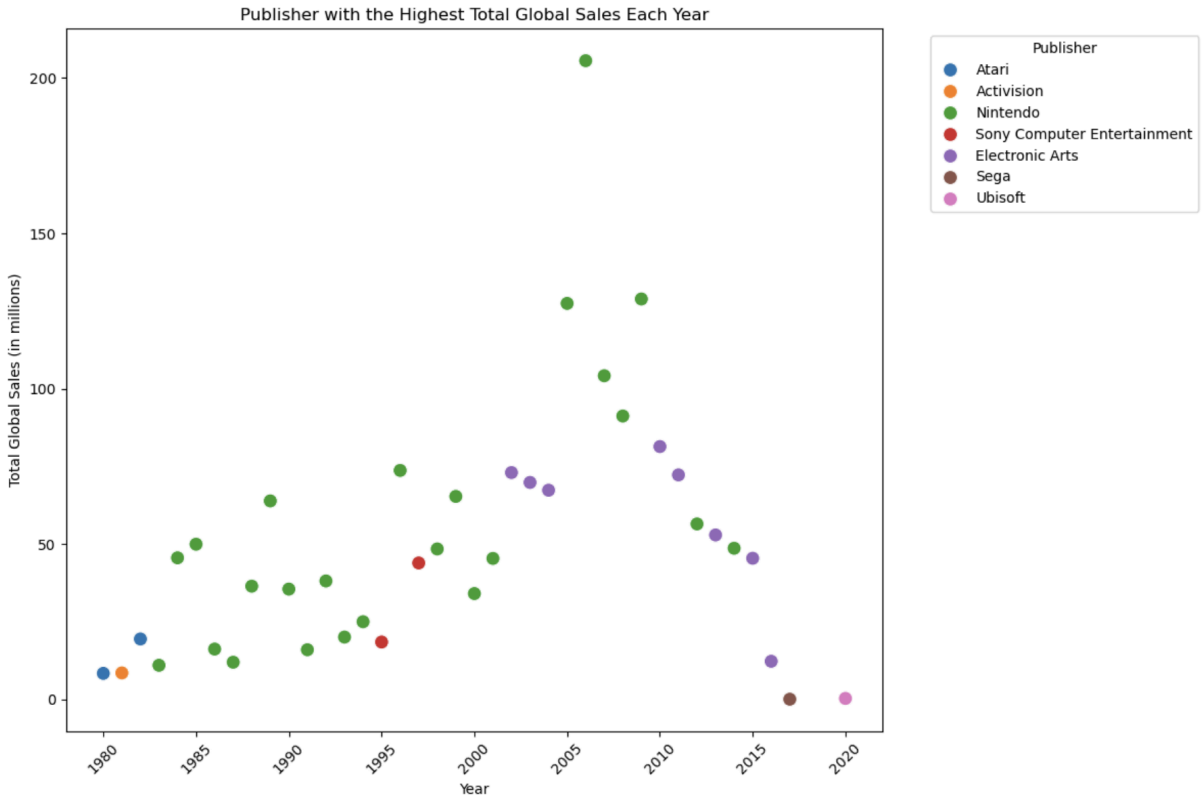
Ria Vellanki



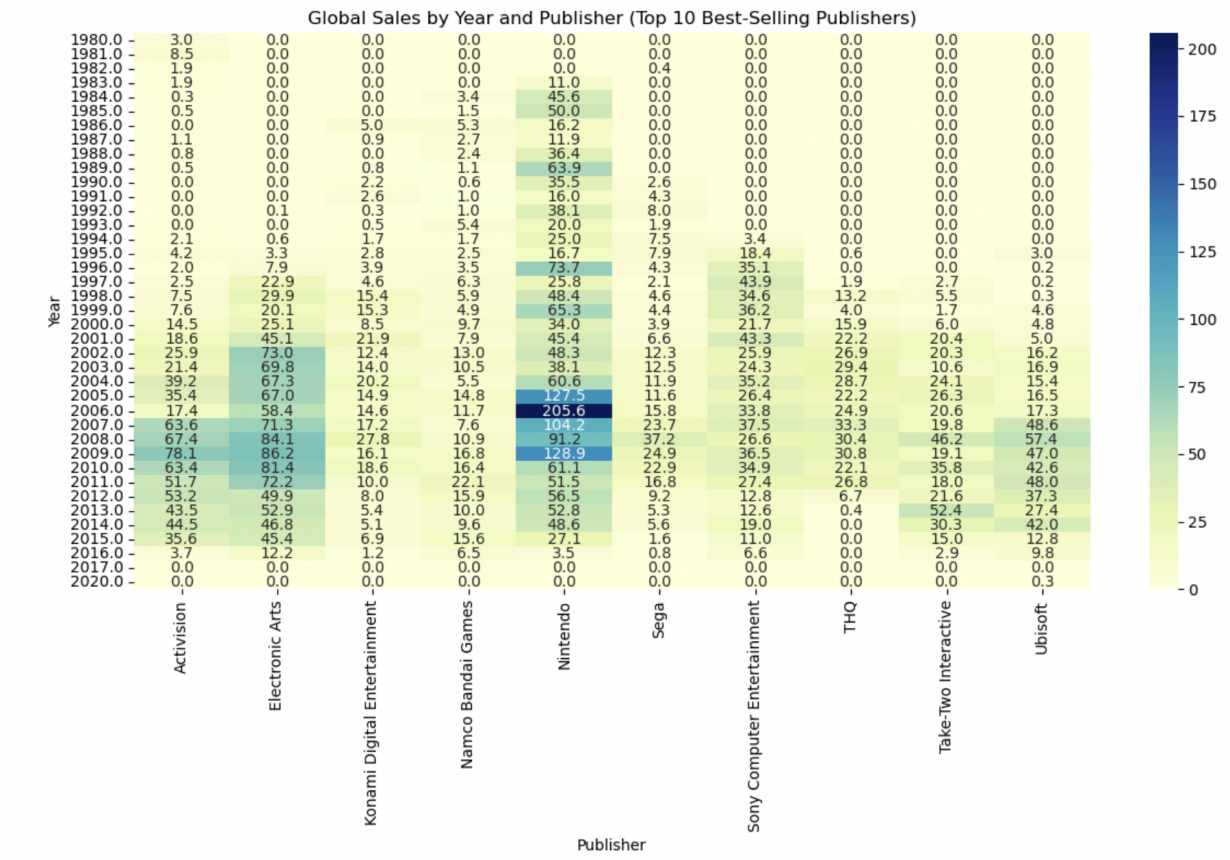
This graph shows that Electronic Arts produced 8.2% of all the games in the dataset, the most out of any other publisher. Out of 16291 rows, about 1336 games were produced by Electronic Arts. This number is quite significant compared to the second publisher with the most games, Activision, which is 5.9% of the dataset or 961 games. This is a difference of almost 400 games produced from 1980 to 2020. 10 of the publishers in the dataset produced almost half of the games while 566 of all other publishers produced the other half. These 10 publishers produced almost the same amount of games combined as 566 other publishers from 1980-2020. This visualization shows the variety in the video game market as well as which corporations are industry

leaders. I started to wonder, how much revenue is each publisher generating? Does quantity correlate with revenue?

In order to answer the questions the last visualization raised, I created a scatter plot and heat map to visualize publishers with the highest total global sales each year.



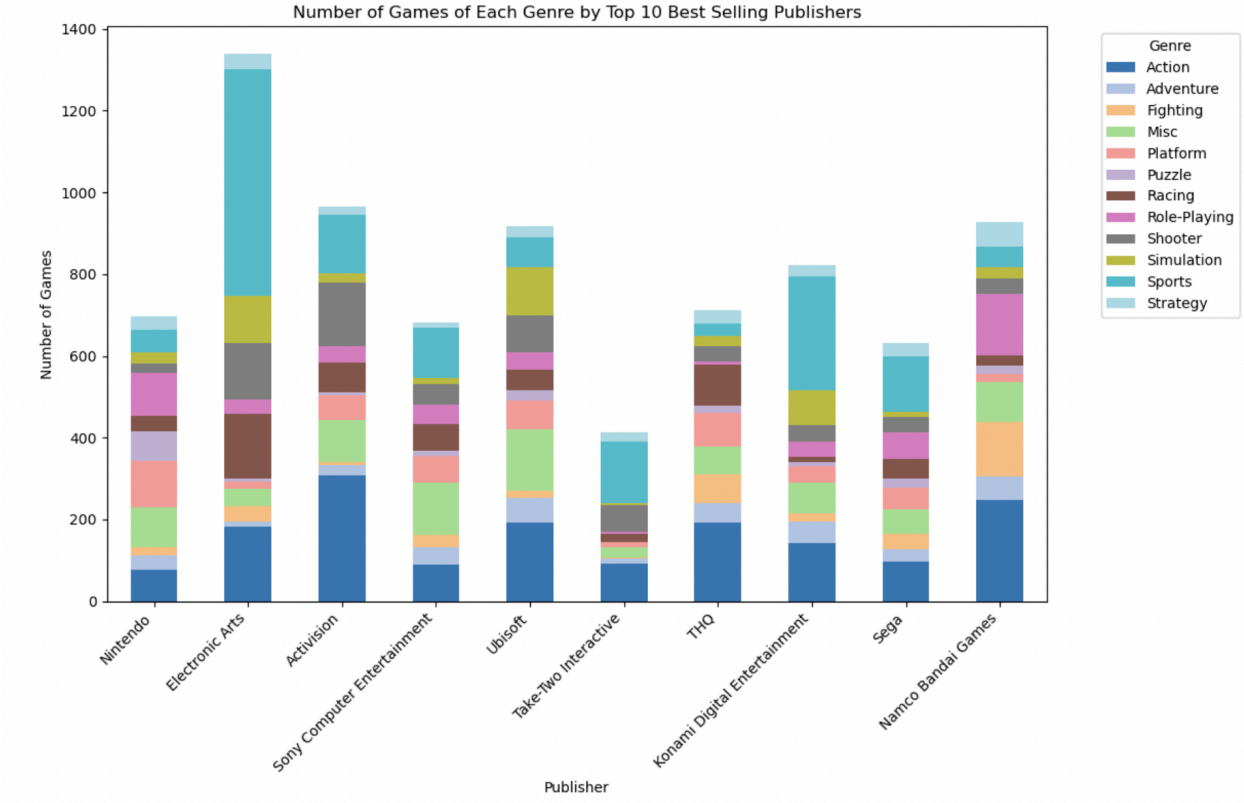
Ria Vellanki



This scatter plot shows that Nintendo is consistently making the most revenue for the majority of the years from 1983-2010. They made more than 200 million from the games they released in 2006. This was quite a surprise because Nintendo only makes up 4.2% of games, which is about half the amount of games produced by Electronic Arts. Electronic Arts has also made significant revenue, the highest being about 86 million in 2009. The difference between the most revenue generated in a year is staggering as Electronic Arts made less than half of Nintendo's maximum revenue. The results of this visualization led me to think about what would warrant Nintendo to generate that much money. I started to wonder whether the genre of the games produced by these publishers could explain the generated revenue.

Ria Vellanki

The next visualization is a stacked bar chart that visualizes the number of games of each genre by top 10 best selling publishers. In this visualization, I am focusing on the top 10 publishers that generated the most revenue and what genre of games they produce.



When looking at this visualization, I noticed that the most produced genres are sports, action, and miscellaneous across all of the publishers. Electronic Arts heavily focuses on sports, action, racing, and shooter whereas Nintendo seems to have a more even range of all genres but heavily focus on platform, role-play, miscellaneous, and action games. This visualization gives insight into why Nintendo has repeatedly found success. Although they may not produce the most games, Nintendo produced games that are easy to pick up with a limited background in video games which allows them to appeal to a wider audience. Platform games are simple games where the main objective is

to move the character through different points in the game. Role-playing games are also relatively easy to pick up as the game is guided by the predetermined story line. This differs from Electronic Art, which heavily produces sports, shooter, action, and racing. Most of these genres typically require some skill and background in video games, narrowing the target audience for the publisher.

Throughout this project I have learned that one of the most important factors for success as a video game publisher is genre and appealing to a wide audience in order to maximize earning potential. Creating beginner friendly and engaging games attracts people from many different gaming backgrounds. Most Nintendo games such as Wii Sports, Duck Hunt, and Legend of Zelda are beginner friendly and perfect for kids and adults. Due to Nintendo being a large corporation that has been in the gaming industry for decades, they have the resources to build engaging backstories and worlds with lore to keep users entertained by games that are quite simple at their core. Nintendo's focus on game genres that are beginner friendly and ability to create exciting worlds with their extensive resources contributed to their ability to generate very large amounts of revenue.

Impact

Working with data has many ethical impacts depending on the type of data used. Video game sales data is relatively safe, as the data is not personal in most manners. This project can help businesses understand how to get to a point of success similar to Nintendo. Although this project focuses on video games, the underlying themes are that in order to make money successfully it is important for a business to create a quality product that appeals to a large audience. Some perspective or data missing from this

project could be price. Video games in general are expensive and not everyone can afford them. There is a possibility that expensive video games tend to get less sales since not as many people would be able to afford them. Alternatively, cheaper and affordable games of high quality have a chance of being bought by people of all economic backgrounds.