Business Intelligence | MIS 5342

Video Game Sales Analysis

Dallas Cowboys | May 4, 2023

Table of Contents

1. Introduction 3
1.1 Objective 3
2. Data 4
2.1 Data Source 4
2.2 Data Attributes 4
2.3 Data Processing and Engineering 5
2.4 Limitations 6
3. Analyses & Findings 6
3.1 Sales Performance by Genre 6
3.2 Sales Performance by Publisher 8
3.3 Meta-Score Distribution 9
3.4 User Rating Distribution 10
3.5 Decision Tree Analysis 10
4. Implications 14
5. Conclusion 15
6. References 15
Appendix 16

Table of Figures

Figure-1: Total Games Sales by Genre 6
Figure-2: Average Global Sales by Genre 7
Figure-3: Total Game Sales by Publisher 8
Figure-4: Average Game Sales by Publisher 8
Figure-5: Distribution of Meta Scores 9
Figure-6: Distribution of User Rating 10
Figure-7: Trial 0 of the Decision Tree and Evaluation Summary 12
Figure 8: Unweighted Confusion Matrix 13
Figure 9: Weighted Confusion Matrix 13

1. Introduction

The video game industry has experienced significant growth over the past decade. The number of players and revenue have increased tremendously, and even during the pandemic, the industry continued to grow unlike any other. According to the World Economic Forum, the gaming industry is projected to maintain its growth trajectory, with a market size expected to reach $321 billion by 2026.

In our video game sales analysis, we will utilize a video game dataset to identify interesting trends in the industry over the past 25 years. We will observe sales patterns by video game genre and publisher to gain insight into how these factors impact revenue generation. Additionally, we will analyze the distribution of meta-scores and user ratings to uncover potential insights. We will then use independent variables such as genre, publisher, meta-score, and user rating to predict the sales performance of each game across the globe. This will allow us to understand where the industry is headed and what characteristics a game must possess to generate more revenue and appeal to a larger audience.

1.1 Objective

As consultants for the video game industry, our objective is to help businesses answer the following questions through our analysis:

Descriptive questions:

- What are the best-selling and worst-selling video game genres?
- How are Metacritic scores and user ratings distributed across video games over time?
- What is the historical relationship between Metacritic scores, user ratings, and global sales?

Predictive Questions:

- Which publishers sell the most games, and is this due to having more games on the market or other factors?
- What are the most important factors for predicting a game's sales performance?

The questions mentioned above are crucial for us to gain a comprehensive understanding of the gaming industry. The descriptive questions will help us uncover interesting trends and patterns in

the gaming industry at both the regional and global levels. These insights will provide a more detailed view of the industry's characteristics and the direction in which it is heading. On the other hand, the predictive questions will shed light on the key factors that determine a game's success in different regions, allowing gaming companies to optimize their strategies and increase their revenue. Furthermore, companies can use our models and results to estimate the expected sales performance for new games based on various features, such as genre, publisher, meta scores, and user ratings. For potential investors, our answers to these questions can provide valuable information for making informed investment decisions, such as what type of game and which publisher to invest in.

2. Data

2.1 Data Source

The data was sourced from Kaggle. Initially, there were two datasets: one containing sales data by region for video games from 1996 to 2022, and the other containing meta-scores and user ratings for video games during the same period. The two datasets were merged by matching the game titles, resulting in the creation of our original dataset.

2.2 Data Attributes

Our merged dataset includes the name, platform, year, genre, publisher, meta score, user rating, and summary for each video game. Sales data is divided into North America, Europe, Japan, and other countries, with sales units measured in millions of dollars. Our primary objective is to gain insights into the video game industry by examining how game genre, publisher, meta-score, and user ratings affect sales performance.

The dataset comprises 4,790 records and 13 variables, with Name, Platform, Year, Genre, and Publisher being qualitative variables, while NA_Sales, EU_Sales, JP_Sales, Other_Sales, Global_Sales, Meta_Score, and User_Rating are quantitative variables. Our analysis will focus on the Platform, Year, Genre, Publisher, NA_Sales, EU_Sales, JP_Sales, Other_Sales, and Global_Sales variables. By examining the factors that affect video game sales across regions and globally, we will address our business questions and draw implications from our findings. A more detailed breakdown of these variables is shown in Table-1.

Table-1: Description of Variables in the Original Dataset

Variable	Description
Name	This is the name of the video game.
Platform	This is the platform or game console that the video game is played on.
Year	This is the year when the video game is published.
Genre	This is the video game genre such as Sports, Racing, Role-Playing, etc.
Publisher	This is the publisher of video games such as Nintendo, Activision, Ubisoft, etc.
NA_Sales	This is the sales data of the video game in North America in millions of USD.
EU_Sales	This is the sales data of the video game in Europe in millions of USD.
JP_Sales	This is the sales data of the video game in Japan in millions of USD.
Other_Sales	This is the sales data of the video game in other countries in millions of USD.
Global_Sales	This is the sales data of the video game in the world in millions of USD.
Meta_Score	This is the score for the video game from Metacritic.
User_Rating	This is the rating for the video game from players and users across the world.
Summary	This is the summary of the game and what the game is about.

2.3 Data Processing and Engineering

We excluded unnecessary variables such as Name and Platform from our analysis. We replaced a few 'tbd' values in the *User_Rating* field with the *Meta_Score* value divided by 10, as the *Meta_Score* was on a scale of 100 and the *User_Rating* was on a scale of 10. We added log sales columns for regional and global sales by taking the logarithm of the respective variables, normalizing the data for conducting various statistical analyses. We also added an Age column, calculating years passed since a video game's release by subtracting the release year from the current year.

We transformed categorical variables such as Genre and Publisher into factor variables for our decision tree analysis The major transformations also included performing a log calculation on all numeric columns to explore and analyze the logarithmic distributions of the variables as well as the standard distributions. Additional variables were created for chi-squared and t-test statistical analysis, creating a binary variable for the predicted top and bottom three genres and for the top and bottom game developers by number of games in the dataset. We ended up with a total of 31 variables, of which we used 25 for various exploratory visualization and analyses.

2.4 Limitations

There are several limitations to our dataset that must be considered:

1. The dataset's sales data is limited to four regions, North America, Europe, Japan, and other countries, which may not represent the entire world and may lead to biased findings.
2. Our dataset lacks variables that consider the technological advancements in video game development over the past 25 years, which could significantly impact sales performance.
3. The dataset solely contains information about released video games, and no information about cancelled or unreleased games, which may constrain the scope of our analysis and prevent examination of factors contributing to game cancellations.
4. Qualitative variables such as player feedback and reviews, which can provide valuable insights into the factors that influence sales performance, are not available in our dataset.
5. Confounding variables that were not considered in our analysis may affect the relationship between variables. For example, cultural factors may influence the popularity of certain game genres and impact sales performance.

3. Analyses & Findings

3.1 Sales Performance by Genre

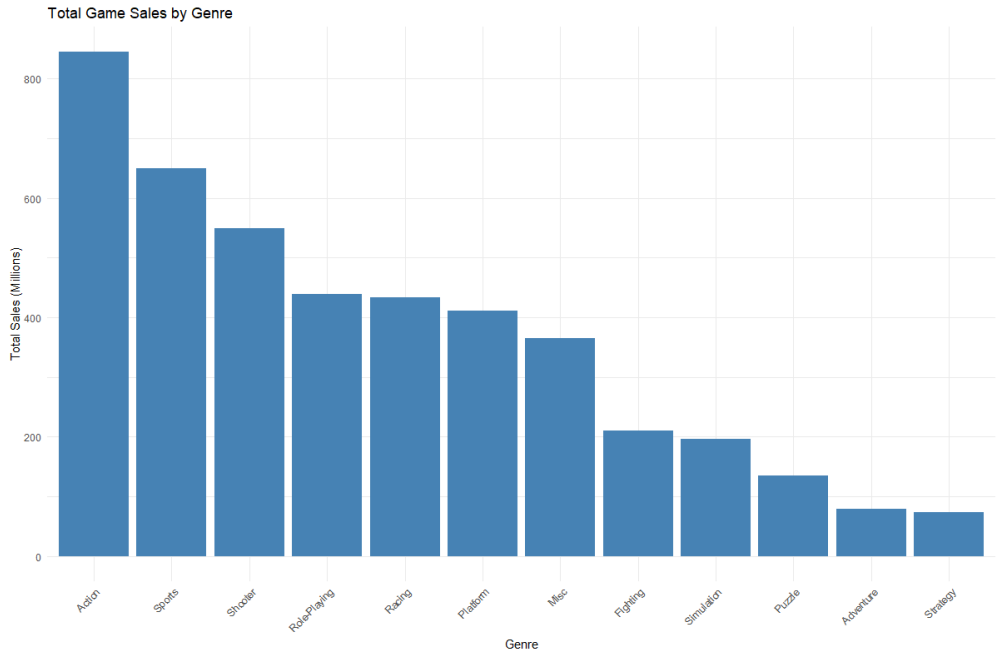

Figure-1: Total Games Sales by Genre

Figure-1 displays the global total game sales by genre. The chart reveals that Action and Shooter games have the highest total sales, which is not surprising given that these genres are currently the most popular in the market. On the other hand, Adventure and Strategy games have the lowest total sales, primarily because they tend to be less expensive, and some are even available for free. Consequently, gamers are less likely to purchase them.

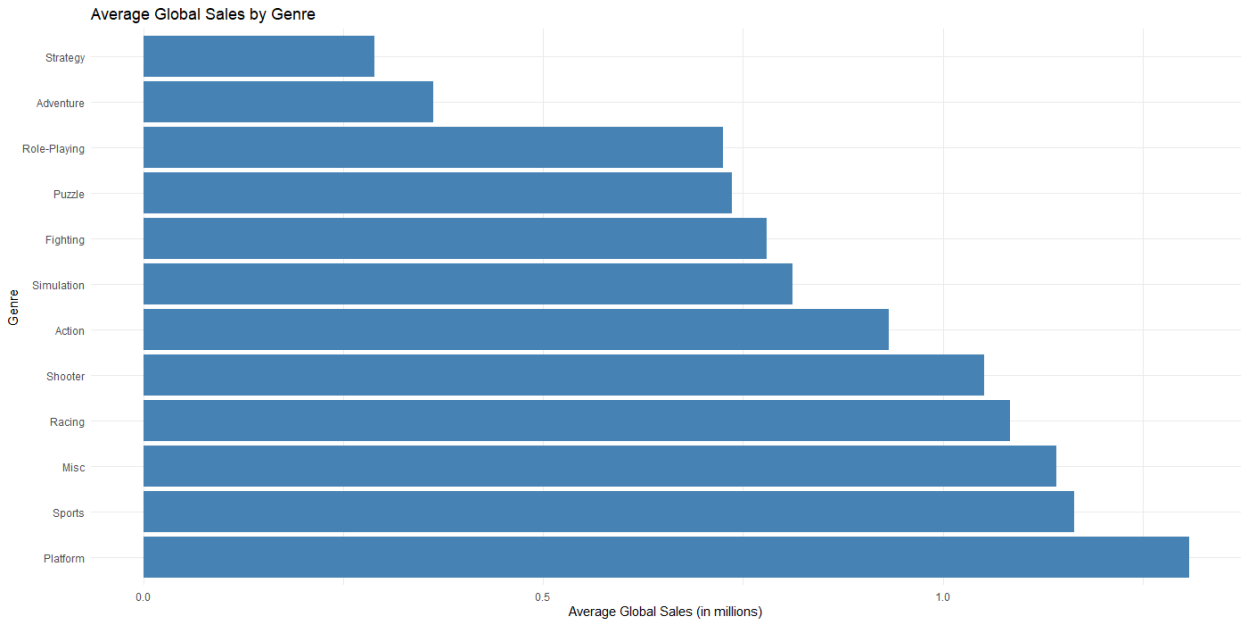

Figure-2: Average Global Sales by Genre

Figure-2 illustrates the average global sales by genre. Interestingly, Platform and Sports have emerged as the top-performing genres with the highest average sales. Conversely, the two genres with the lowest average sales are Strategy and Adventure, which is consistent with their low total game sales.

3.2 Sales Performance by Publisher

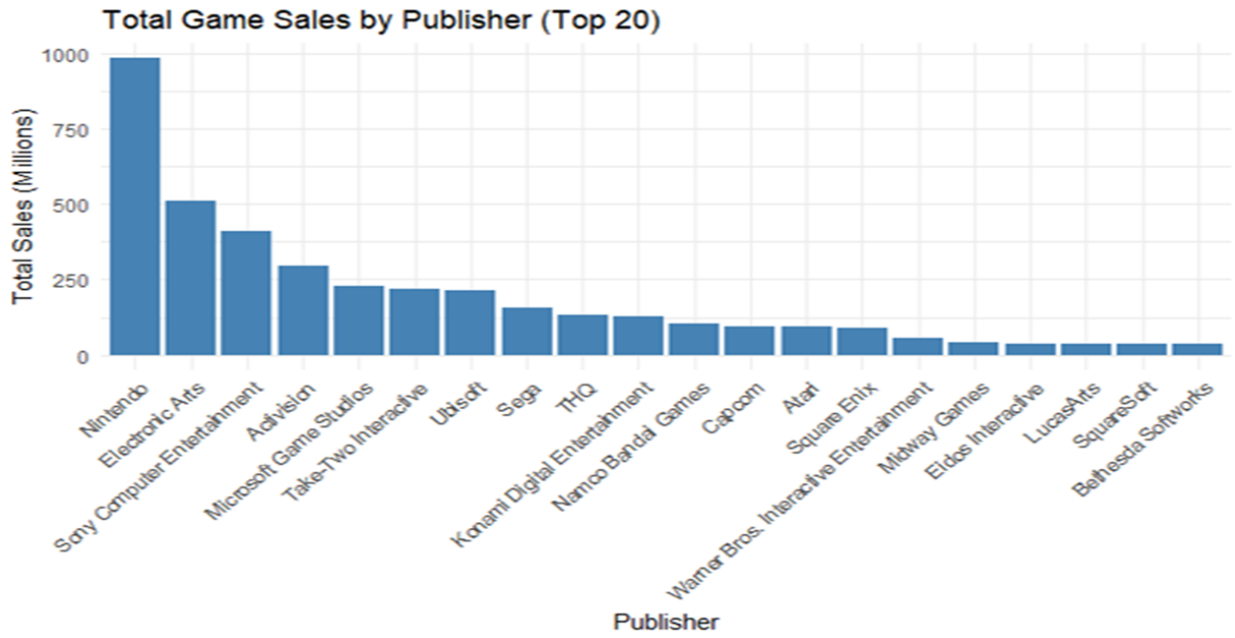

Figure-3: Total Game Sales by Publisher

Figure-3 shows total game sales of the top 20 publishers where Nintendo, Electronic Arts and Sony Computer Entertainment ranked in the top 3, while Lucas Arts, Square Soft, and Bethesda Softworks ranked in the bottom 3.

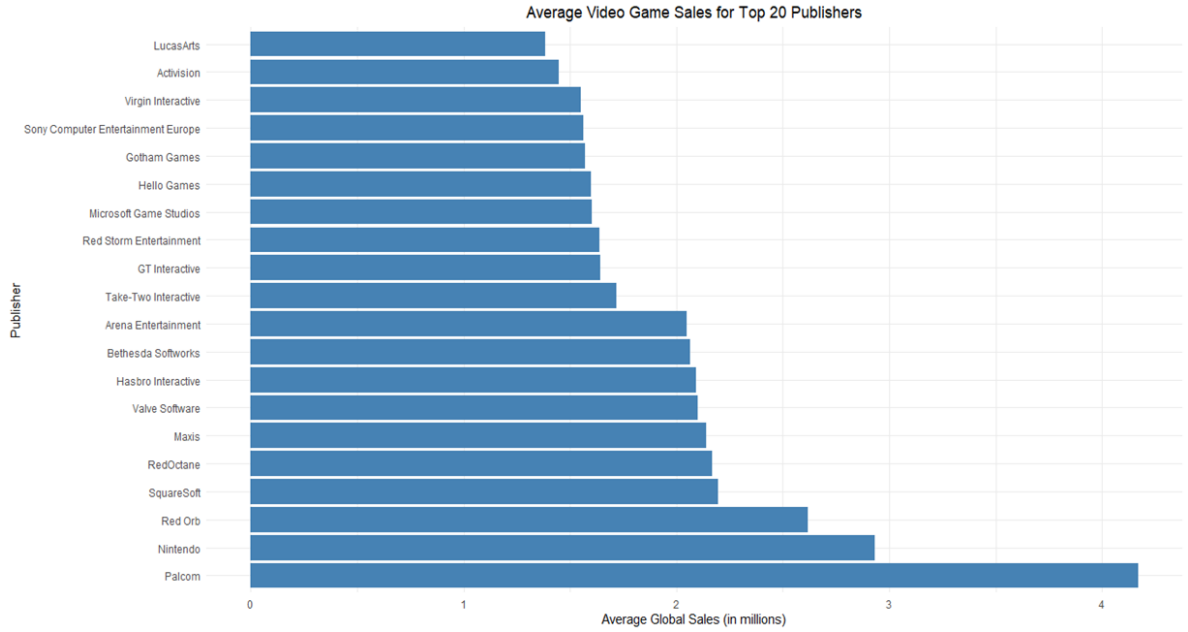

Figure-4: Average Game Sales by Publisher

In Figure-4, we have plotted the average global sales against publishers. Surprisingly, Palcom is leading the chart with the highest average global sales, with Nintendo falling behind at the second position. It is worth noting that Palcom and Nintendo were direct competitors in the gaming industry for a few years. On the other end of the spectrum, Activision's ranking has dropped significantly from the total sales chart, and now LucasArts has the lowest average video game sales for the top 20 publishers.

3.3 Meta-Score Distribution

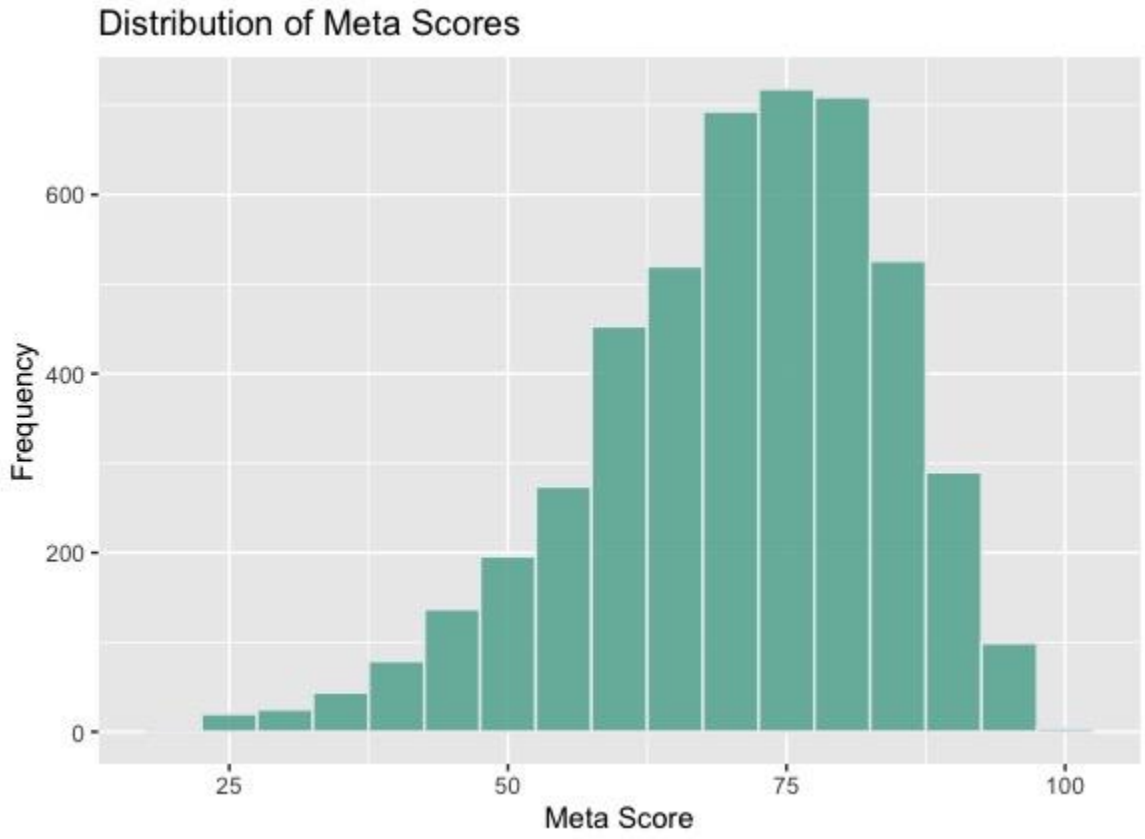

Figure-5: Distribution of Meta Scores

In Figure-5, which depicts a histogram illustrating the distribution of meta scores on a scale of 0 to 100. The data distribution is skewed to the left, indicating that the majority of the meta scores are in the higher range. The most common meta score range falls between 65 and 85, and over 50% of the games fall within this range.

3.4 User Rating Distribution

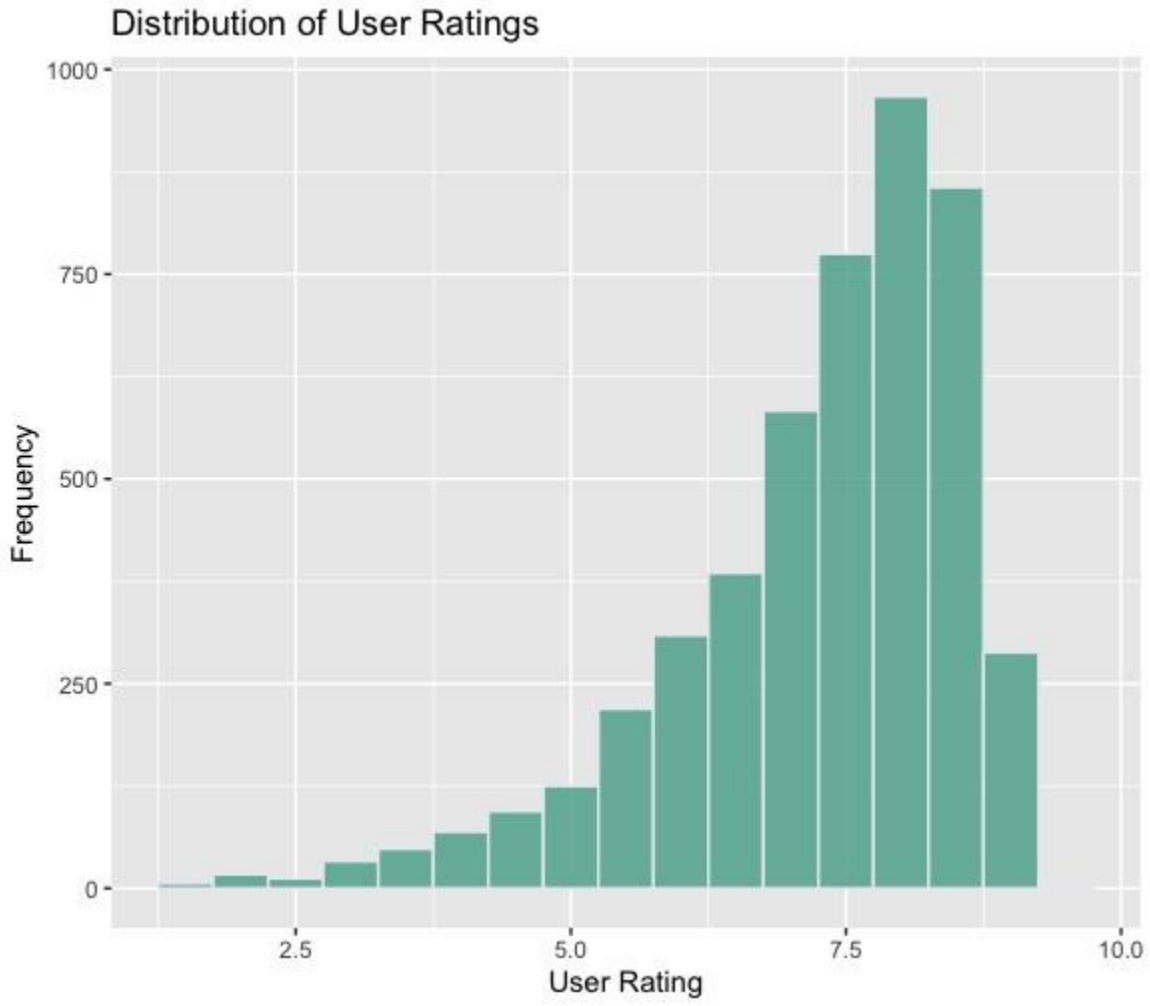

Figure-6: Distribution of User Rating

Figure-6 displays a histogram representing the distribution of user ratings on a scale of 0 to 10. The histogram is skewed to the left, indicating that user ratings are normally on the higher end. The typical user rating falls between 7 and 9, which translates to a range of 70 to 90 when converted to a scale of 100. When compared to the distribution of meta scores in Figure-5, the distribution of user ratings tends to be higher overall. This may be due to the fact that users are generally more subjective and less critical than expert critics when it comes to evaluating video games.

3.5 Decision Tree Analysis

We wanted to test the correlation between our various attributes, namely how the publisher, genre, meta score, and user score variables related to regional and global total and average sales. To do

this, we performed correlation tests on the continuous variables and chi-squared tests on the categorical variables. Below is a table with the relevant tests performed, the test statistic result, and the p-value associated with the test statistic for our data set.

Table-2: Correlation Tests

Test Name	**Test Statistic**	**P-value**
User Rating vs NA Sales (correlation)	t = 0.1135	3.419e-15
User Rating vs EU Sales (correlation)	t = 0.0834	7.589e-09
User Rating vs JP Sales (correlation)	t = 0.1628	< 2.2e-16
User Rating vs Other Sales (correlation)	t = 0.0820	1.351e-08
Meta Score vs NA Sales (correlation)	t = 0.2442	< 2.2e-16
Meta Score vs EU Sales (correlation)	t = 0.2208	< 2.2e-16
Meta Score vs JP Sales (correlation)	t = 0.1858	< 2.2e-16
Meta Score vs Other Sales (correlation)	t = 0.2041	< 2.2e-16
Top 5% Publishers - Global Sales (one-sample t-test)	t = 1.222362	< 2.2e-16
Top 5% Publishers - Average Sales (one-sample t-test)	t = 0.08443378	< 2.2e-16
Bottom 5% Publishers - Global Sales (one-sample t-test)	t = 0.3485714	0.04494
Bottom 5% Publishers - Average Sales (one-sample t-test)	t = 0.02076651	0.02637
User Rating vs Total Sales (Pearson's correlation)	t = 14.696	< 2.2e-16
User Rating vs Average Sales (Pearson's correlation)	t = 8.4367	< 2.2e-16
Meta Score vs Total Sales (Pearson's correlation)	t = 31.729	< 2.2e-16
Meta Score vs Average Sales (Pearson's correlation)	t = 27.14	< 2.2e-16
Action, Sports, Shooter Genres vs Total Sales (Pearson's Chi-squared)	X^2 = 13.844	0.003126
Action, Sports, Shooter Genres vs Average Sales (Pearson's Chi-squared)	X^2 = 6.9295	0.07418
Simulation, Adventure, Strategy Genres vs Total Sales (Pearson's Chi-squared)	X^2 = 32.345	4.426e-07
Simulation, Adventure, Strategy Genres vs Average Sales (Pearson's Chi-squared)	X^2 = 12.169	0.006827

From this we gathered the variables that we were going to use for the decision tree analysis. We used the c50 package in R to run the regression and make the model. Our target variable was the Global_sales data column and this variable was converted to a binary “High” / “Low” variable for analysis purposes. The data used for the regression included factor columns for the publisher and genre, where each genre and publisher were represented by an integer factor that has a range of all the unique items in each of the genre and publisher columns.Then we also included the user score and meta score columns, unaltered aside from the cleaning that was done at the start of analysis. From there we iterated through different regressions, starting with one trial and ending at fifty trials to find the optimal number of trials for reducing errors in the confusion matrix. Eleven trials were optimal for the unweighted data set, so we ran that regression, and this is the result of that decision tree summary.

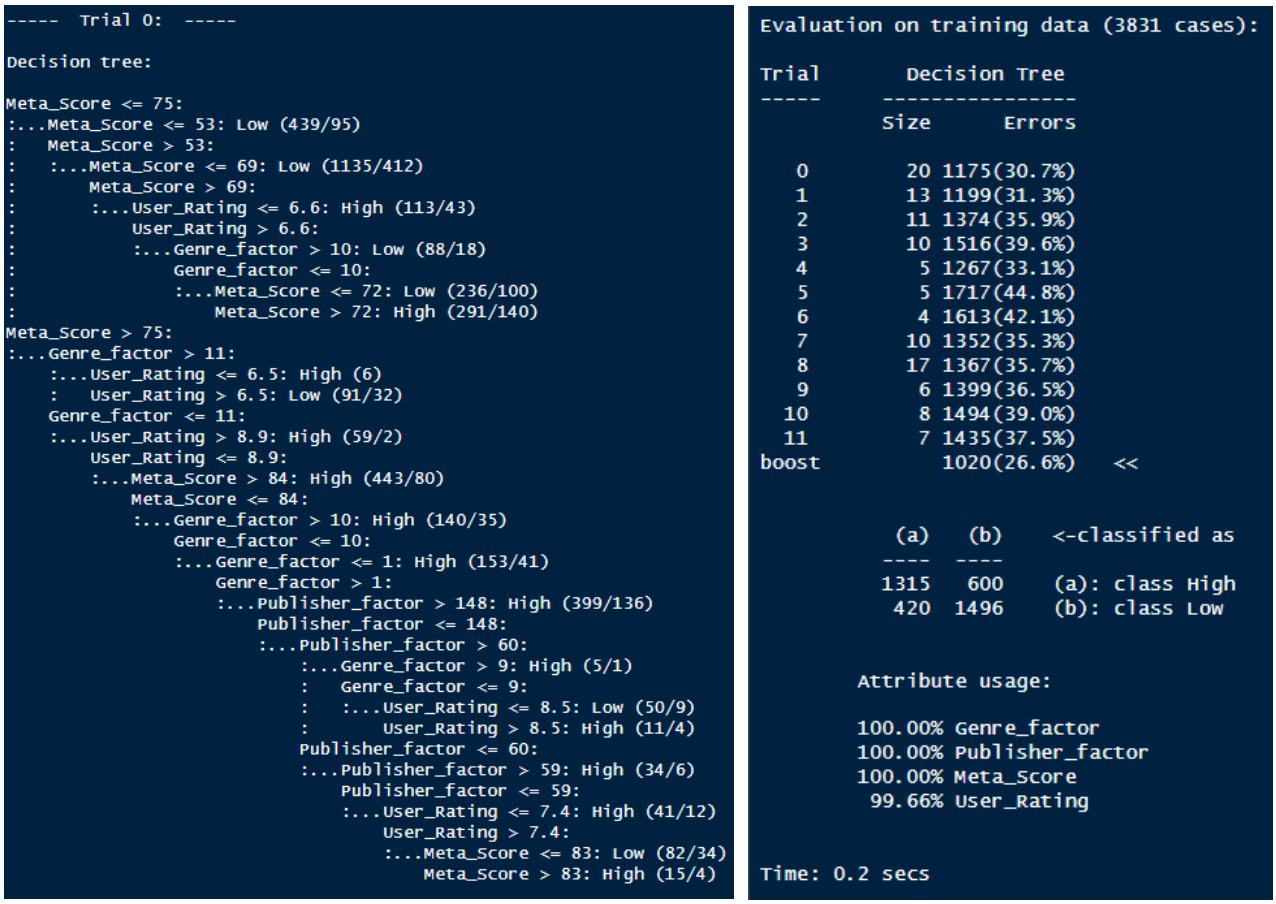

Figure-7: Trial 0 of the Decision Tree and Evaluation Summary

We also created a cost matrix to weigh the regression to try to reduce the amount of "highly undesirable errors" where the model predicts the game hitting a sales target when, in fact, that sales target will not be reached, a Type 2 error. The weights of the cost matrix were zeros for predicting the correct outcome, one for predicting not hitting target when the target was actually reached, and five for predicting hitting the target when the target would not be reached. This was determined to be a fair cost weight, as it prioritizes avoiding type 2 errors while still incentivizing the model to care about avoiding Type 1 errors. We reran the regression and then determined a confusion matrix for both the unweighted and the weighted regression results.

Total Observations in Table: 958

	Predicted		
Actual	High	Low	Row Total
High	308	180	488
	0.753	0.328	
Low	101	369	470
	0.247	0.672	
Column Total	409	549	958
	0.427	0.573	

Figure 8: Unweighted Confusion Matrix

	Predicted		
Actual	High	Low	Row Total
High	216	272	488
	0.797	0.396	
Low	55	415	470
	0.203	0.604	
Column Total	271	687	958
	0.283	0.717	

Figure 9: Weighted Confusion Matrix

4. Implications

After conducting both descriptive and predictive analyses, we have discovered that certain variables have a significant impact on video game sales performance. These variables include game genre, publisher, metacritic score, and user rating. Our decision tree model aims to minimize overall error, resulting in an accuracy of 70.67%, or to minimize the worst errors by half, with an overall accuracy of 65.86%.

Our analysis also reveals some interesting trends. Action and sports games are the most popular genres, with sports games ranking second even when looking at average sales. However, platform games lead the average sales table. Metacritic scores are left-skewed, with a typical score between 65 and 85, as experts tend to be more objective in their reviews. User ratings are also left-skewed, with the typical score between 7 and 9. In terms of publishers, Nintendo and Electronic Arts capture the most sales, but Palcom surpasses both when we look at average sales. Palcom was a subsidiary of Konami, a direct competitor of Nintendo, and was only around for a few years, which led to its high average units sold.

Based on the analysis, video game developers should consider focusing on developing games in the action and sports genres since they have the highest total sales. They should also aim to produce high-quality games that receive positive Metacritic scores and user ratings, as these variables have a significant impact on sales performance.

Developers should also pay attention to the publisher they choose to work with, as the publisher's reputation and marketing strategies can affect the game's sales. Nintendo and Electronic Arts are the top publishers in terms of total sales, but Palcom has the highest average sales, possibly due to its short lifespan and unique position in the market.

Furthermore, developers should consider offering free or low-cost versions of their adventure and strategy games, as these genres tend to have lower sales due to their pricing and competition from free games.

Overall, the analysis suggests that developers should focus on developing high-quality games in popular genres while also considering the publisher they work with and pricing strategies.

5. Conclusion

The analysis of video game sales provides valuable insights for various stakeholders in the gaming industry, such as business managers, investors, game developers, publishers, and players. It enables business managers to gain a comprehensive understanding of the industry and its major players. Investors can use the analysis to identify profitable investment opportunities in the market. Game developers seeking to achieve high levels of gameplay and success can make informed decisions on which publishers best serve their goals based on the analysis. Publishers equipped with this knowledge have better market insights, such as which game genres are trending and where to allocate resources for maximum profit.

Additionally, the analysis can be beneficial to game players, providing them with information on the top trending games in the market and how their favorite games compare to others. Our analysis provides concrete data that all stakeholders can use to make informed business decisions, such as determining the best regions to prioritize sales, identifying game development areas to cut losses, targeting markets for business expansion, or choosing a gaming company to acquire or merge with for competitive advantage.

In conclusion, our data analysis offers highly sought-after business intelligence insights for all stakeholders in the gaming industry to make better decisions.

6. References

1. "Gaming Is Booming and Is Expected to Keep Growing. This Chart Tells You All You Need to Know." *World Economic Forum*, 28 July 2022, https://www.weforum.org/agenda/2022/07/gaming-pandemic-lockdowns-pwc-growth/.
2. Top Video Games 1995-2021 Metacritic, https://www.kaggle.com/datasets/deepcontractor/top-video-games-19952021-metacritic
3. Smith, Gregory. *Video Game Sales*. https://www.kaggle.com/datasets/gregorut/videogamesales. Accessed 4 May 2023.

Appendix

Decision Tree Trials

Trial 1

```
-----  Trial 1:  -----

Decision tree:

Meta_Score > 78: High (1119.7/361.1)
Meta_Score <= 78:
:...Meta_Score <= 53: Low (420.5/123.1)
    Meta_Score > 53:
    :...Publisher_factor <= 7: Low (58.8/12.5)
        Publisher_factor > 7:
        :...Publisher_factor <= 9: High (101.1/33.3)
            Publisher_factor > 9:
            :...Genre_factor <= 5:
                :...Publisher_factor > 134: High (601/264)
                :   Publisher_factor <= 134:
                :   :...Meta_Score <= 77: Low (308.1/122.7)
                :       Meta_Score > 77: High (22.5/5.6)
                Genre_factor > 5:
                :...Genre_factor <= 6: Low (95.1/25.9)
                    Genre_factor > 6:
                    :...Publisher_factor > 222: Low (231.2/68.3)
                        Publisher_factor <= 222:
                        :...Publisher_factor > 216: High (45.4/4.3)
                            Publisher_factor <= 216:
                            :...Publisher_factor <= 59: Low (165.9/57.4)
                                Publisher_factor > 59:
                                :...Publisher_factor <= 62: High (129.6/41.9)
                                    Publisher_factor > 62: Low (532.3/230.7)
```

Trail 2

```
-----  Trial 2:  -----

Decision tree:

Meta_Score > 86: High (331.4/86.6)
Meta_Score <= 86:
:...Genre_factor > 11: Low (159.9/47.6)
    Genre_factor <= 11:
    :...Publisher_factor > 264: Low (27.7/3.7)
        Publisher_factor <= 264:
        :...Publisher_factor > 262: High (33.5/4.4)
            Publisher_factor <= 262:
            :...Meta_Score <= 62: Low (886.2/359.3)
                Meta_Score > 62:
                :...Publisher_factor <= 5: Low (33.2/5.4)
                    Publisher_factor > 5:
                    :...Publisher_factor > 222: Low (470.7/211.2)
                        Publisher_factor <= 222:
                        :...Publisher_factor > 212: High (84.1/19.7)
                            Publisher_factor <= 212:
                            :...Genre_factor > 8: High (550.3/220.9)
                                Genre_factor <= 8:
                                :...Meta_Score <= 84: Low (1186/587.5)
                                    Meta_Score > 84: High (68.1/21.4)
```

Trail 3

```
-----  Trial 3:  -----

Decision tree:

Genre_factor > 11: Low (190.2/68.2)
Genre_factor <= 11:
:...Publisher_factor <= 7: Low (93.9/29.6)
    Publisher_factor > 7:
    :...Meta_Score > 89: High (148.5/29.7)
        Meta_Score <= 89:
        :...Genre_factor <= 1: High (709.9/309)
            Genre_factor > 1:
            :...Genre_factor <= 2: Low (163.4/52.2)
                Genre_factor > 2:
                :...Publisher_factor > 173:
                    :...Publisher_factor <= 197: Low (48.4/12.2)
                    :   Publisher_factor > 197: High (961.8/466)
                    Publisher_factor <= 173:
                    :...Publisher_factor > 153: High (358.8/114.4)
                        Publisher_factor <= 153:
                        :...Publisher_factor <= 60: High (697.3/287.7)
                            Publisher_factor > 60: Low (458.8/178.5)
```

Trail 4, 5, and 6

```
-----  Trial 4:  -----

Decision tree:

Publisher_factor > 264: Low (28.8/3.8)
Publisher_factor <= 264:
:...Publisher_factor > 262: High (35.5/5.4)
    Publisher_factor <= 262:
    :...User_Rating > 8.8: High (100.4/26.5)
        User_Rating <= 8.8:
        :...Meta_Score <= 75: Low (2308.6/1033.8)
            Meta_Score > 75: High (1357.6/588.8)

-----  Trial 5:  -----

Decision tree:

Publisher_factor > 264: Low (27.6/4.1)
Publisher_factor <= 264:
:...Meta_Score <= 39: Low (78.6/18.5)
    Meta_Score > 39:
    :...Publisher_factor > 262: High (34.1/5.8)
        Publisher_factor <= 262:
        :...Genre_factor <= 11: High (3506.8/1686.5)
            Genre_factor > 11: Low (183.9/69.3)

-----  Trial 6:  -----

Decision tree:

Meta_Score > 86: High (299.2/102.9)
Meta_Score <= 86:
:...Meta_Score <= 53: Low (409.1/153.6)
    Meta_Score > 53:
    :...User_Rating <= 5.4: High (171.8/65.9)
        User_Rating > 5.4: Low (2950.8/1427.9)
```

Trail 7

```
-----  Trial 7:  -----

Decision tree:

Publisher_factor > 264: Low (26.2/4.3)
Publisher_factor <= 264:
:...Publisher_factor > 262: High (34.6/5.7)
    Publisher_factor <= 262:
    :...User_Rating > 8.8: High (96.4/28.8)
        User_Rating <= 8.8:
        :...Meta_Score <= 62: Low (919.3/400.4)
            Meta_Score > 62:
            :...Publisher_factor <= 5: Low (34.3/6.5)
                Publisher_factor > 5:
                :...Genre_factor > 11: Low (136.5/57.2)
                    Genre_factor <= 11:
                    :...Publisher_factor <= 164: High (1522.3/659.9)
                        Publisher_factor > 164:
                        :...Publisher_factor <= 207: Low (117.5/29.9)
                            Publisher_factor > 207:
                            :...Publisher_factor <= 222: High (433.1/186)
                                Publisher_factor > 222: Low (510.8/243.4)
```

Trail 8

```
-----  Trial 8:  -----

Decision tree:

Meta_Score > 89: High (145.1/45)
Meta_Score <= 89:
:...Meta_Score <= 39: Low (75.8/22.5)
    Meta_Score > 39:
    :...Genre_factor <= 5:
        :...Genre_factor > 2: High (662.2/295.8)
        :   Genre_factor <= 2:
        :   :...Genre_factor <= 1: High (706.3/323.5)
        :       Genre_factor > 1: Low (163.3/56.7)
        Genre_factor > 5:
        :...Publisher_factor > 250: Low (87.2/26.2)
            Publisher_factor <= 250:
            :...Publisher_factor <= 59: Low (442.2/179.9)
                Publisher_factor > 59:
                :...Publisher_factor <= 60: High (204.3/66.4)
                    Publisher_factor > 60:
                    :...Publisher_factor <= 110: Low (107.9/28.7)
                        Publisher_factor > 110:
                        :...Genre_factor > 9:
                            :...Publisher_factor <= 166: High (196.8/86.1)
                            :   Publisher_factor > 166: Low (260.6/108.3)
                            Genre_factor <= 9:
                            :...Publisher_factor <= 148: Low (82.1/28.8)
                                Publisher_factor > 148:
                                :...Genre_factor > 8: High (188.5/70.9)
                                    Genre_factor <= 8:
                                    :...Publisher_factor > 222: Low (123/51.3)
                                        Publisher_factor <= 222:
                                        :...Publisher_factor > 213: High (55.3/12.2)
                                            Publisher_factor <= 213: [S1]

SubTree [S1]

Publisher_factor <= 164: High (171.3/66.2)
Publisher_factor > 164: Low (159.3/64.7)
```

Trail 9 and 10

```
-----  Trial 9:  -----

Decision tree:

Meta_Score <= 79: Low (2883/1358.3)
Meta_Score > 79:
:...User_Rating <= 6.7: High (53/6.2)
    User_Rating > 6.7:
    :...Genre_factor <= 1: High (126.5/42.1)
        Genre_factor > 1:
        :...Genre_factor > 10: High (169.7/72.8)
            Genre_factor <= 10:
            :...User_Rating <= 7.8: Low (150.4/60)
                User_Rating > 7.8: High (448.4/210.6)

-----  Trial 10:  -----

Decision tree:

Publisher_factor > 161:
:...Publisher_factor <= 164: High (252.4/53.2)
:   Publisher_factor > 164:
:   :...Publisher_factor <= 193: Low (120/33.5)
:       Publisher_factor > 193: High (1481.7/720)
Publisher_factor <= 161:
:...Publisher_factor > 60: Low (836.3/339.7)
    Publisher_factor <= 60:
    :...Publisher_factor > 59: High (279.2/91.9)
        Publisher_factor <= 59:
        :...Publisher_factor > 55: Low (74.8/19.9)
            Publisher_factor <= 55:
            :...Publisher_factor <= 54: Low (744.3/363.2)
                Publisher_factor > 54: High (42.2/11)
```

Trail 11 and 12

```
-----  Trial 11:  -----

Decision tree:

Meta_Score > 84: High (427.5/178.8)
Meta_Score <= 84:
:...Publisher_factor > 164: Low (1434.2/651.1)
    Publisher_factor <= 164:
    :...Publisher_factor > 163: High (196.8/46.7)
        Publisher_factor <= 163:
        :...Publisher_factor > 60: Low (754.5/327.6)
            Publisher_factor <= 60:
            :...Publisher_factor > 59: High (226.6/83.5)
                Publisher_factor <= 59:
                :...Publisher_factor <= 55: High (726.7/342.5)
                    Publisher_factor > 55: Low (64.7/18.5)

-----  Trial 12:  -----

Decision tree:
 High (3817/1867.3)

*** boosting reduced to 12 trials since last classifier is very inaccurate
```