



# Data Science Project Report

## Video Games Data Analysis

### Team 18

Name	Sec	BN
Gheith Omar	2	8
Mohamed Akram	2	12
Mariam Muhammed	2	22
Marim Naser	2	23

Supervised by  
Prof. Dina Elreedy  
Eng. Mohamed Abdallah

# Table of Content

1. [Problem Definition](#)
2. [Motivation](#)
3. [Data Exploration](#)
4. [Data Preparation](#)
5. [Analysis:](#)
  - 5.1. [Descriptive Questions](#)  
[D1](#) [D2](#)
  - 5.2. [Exploratory Questions](#)  
[E1](#) [E2](#) [E3](#) [E4](#) [E5](#) [E6](#)
  - 5.3. [Predictive Questions](#)  
[P1](#) [P2](#) [P3](#)
  - 5.4 [Mechanistic Questions](#)  
[M1](#)
6. [Conclusion](#)

# 1. Problem Definition

The problem at hand is to analyze and understand the global sales of video games across different genres and identify trends in consumer behavior. The goal of this analysis is to provide insights that can inform business decisions in the video game industry and help companies optimize their sales strategies.

To achieve this goal, we need to gather and analyze data on video game sales, including information on the genre, platform, publisher, developer, year of release, ratings, and region of sale.

Our analysis will involve using statistical models and machine learning algorithms to make predictions about future sales. We will also create visualizations to help us better understand the data and communicate our findings to stakeholders.

The ultimate objective of this analysis is to provide actionable insights that can inform marketing, product development, and distribution strategies for video game companies. By understanding consumer behavior and trends in video game sales, companies can better position themselves in the market and maximize their revenue potential.

## 2. Motivation

The motivation for analyzing video game sales data is to uncover valuable insights that can help businesses in the video game industry stay ahead of the competition and succeed in an increasingly crowded market.

By analyzing sales data across different genres and regions, we can identify trends in consumer behavior and preferences, as well as emerging opportunities in the market. This information can inform critical business decisions such as marketing strategies, product development, and distribution channels.

Moreover, the video game industry is a rapidly evolving landscape, with new technologies and platforms emerging on a regular basis. By analyzing sales data and identifying emerging trends, we can help businesses stay on the cutting edge of innovation and adapt to changing market conditions.

Ultimately, the motivation for analyzing video game sales data is to provide businesses with the insights they need to make informed decisions, stay competitive, and maximize their revenue potential. As data scientists, we are committed to leveraging the power of data to drive growth, innovation, and success in the video game industry and beyond.

### 3. Data Exploration

- Our dataset has 16719 rows and 16 columns.
- We have 10 continuous features and 6 discrete features
  - Continuous:
    - **Year\_of\_Release**
    - **NA\_Sales**: North America Sales
    - **EU\_Sales**: Europe Sales
    - **JP\_Sales**: Japan Sales
    - **Other\_Sales**
    - **Global\_Sales**
    - **Critic\_Score**: Aggregate score compiled by Metacritic staff.
    - **Critic\_Count**: The number of critics used in coming up with the Critic\_score.
    - **User\_Score**: Score by Metacritic's subscribers
    - **User\_Count**: Number of users who gave the user\_score
  - Discrete:
    - **Name**
    - **Developer**
    - **Publisher**
    - **Platform**
    - **Genre**
    - **Rating**: ESRB ratings provide information about what's in a game or app so parents and consumers can make informed choices about which games are right for their family, 4 main categories are E (everyone), M(mature +17), T(teenagers) and E10+ (everyone +10).

- Number of unique values in each column:

Year_of_Release	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales	Critic_Score	Critic_Count
39	402	307	244	155	629	82	106

Name	Developer	Publisher	Platform	Genre	Rating	User_Score	User_Count
11562	1696	581	31	12	8	96	888

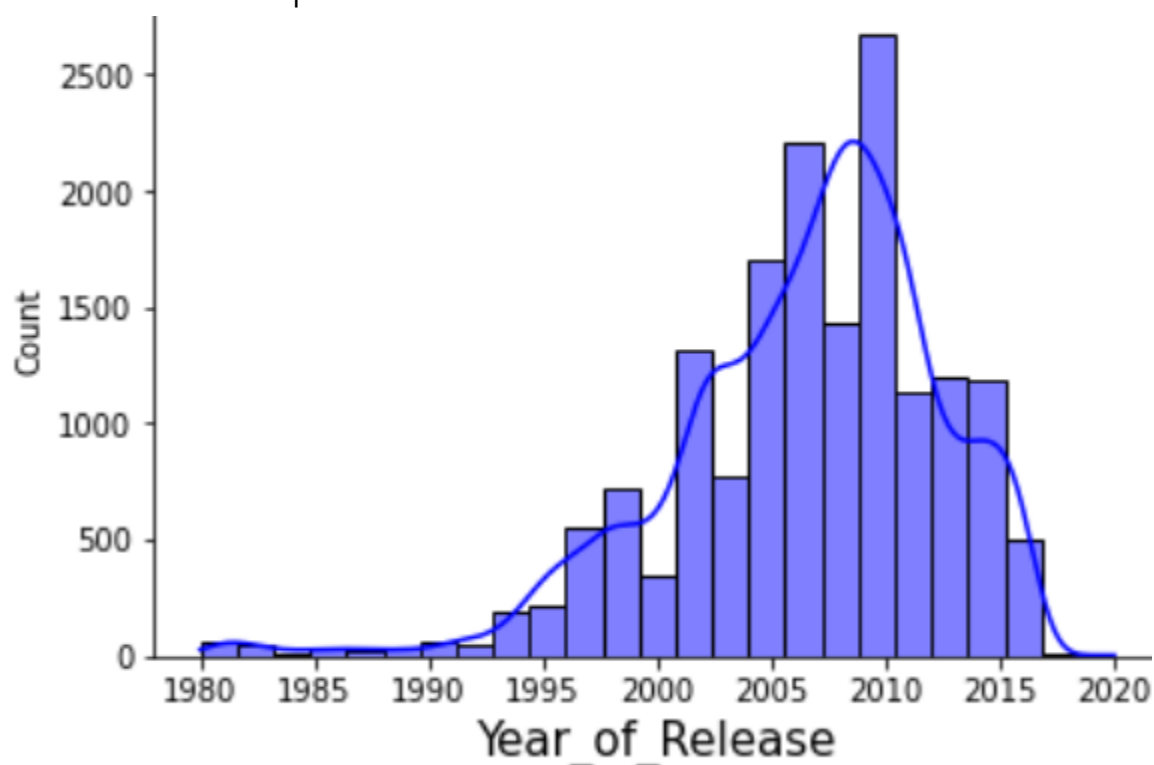
- Missing values:
  - Total Number of rows : 10031
  - Number of rows with missing values: 5954
  - Number of rows after dropping the missing values: 4077
    - rows with missing values are very large > half size of the data, so we can't drop them all. We need to use methods for imputation according to the features used.
  - Missing Values ratios for each column:

User_Count	Critic_Score	Critic_Count	Rating	User_Score	Developer	Year_of_Release	Publisher	Name	Genre
54.6%	51.3%	51.3%	40.5%	40.1%	39.6%	1.6%	0.3%	0.01%	0.01%

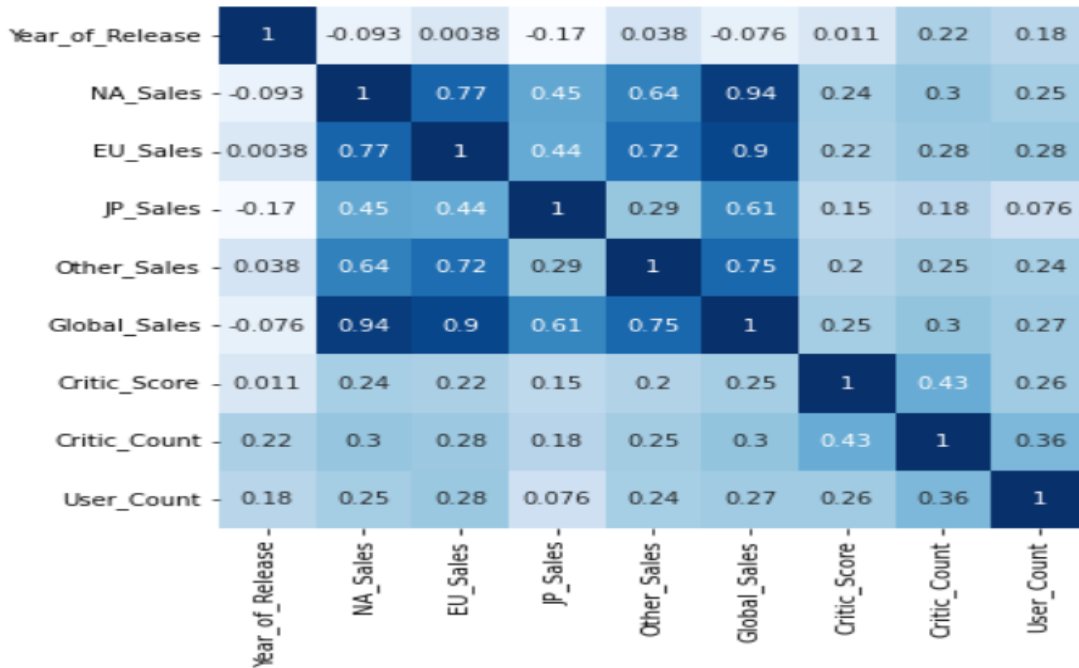
- Sadly critic score, user score, developer, and rating have a very high ratio of missing value however they're key features for any video game so we can't just drop them or even impute with median. We need to try an effective imputation method as KNN.
- Outliers:
  - If we remove all the outliers, we'll have 9470 rows.

Year_of_Release	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales	Critic_Score	Critic_Count	User_Count	All
305	1687	2061	2434	2243	1892	83	251	1080	12036

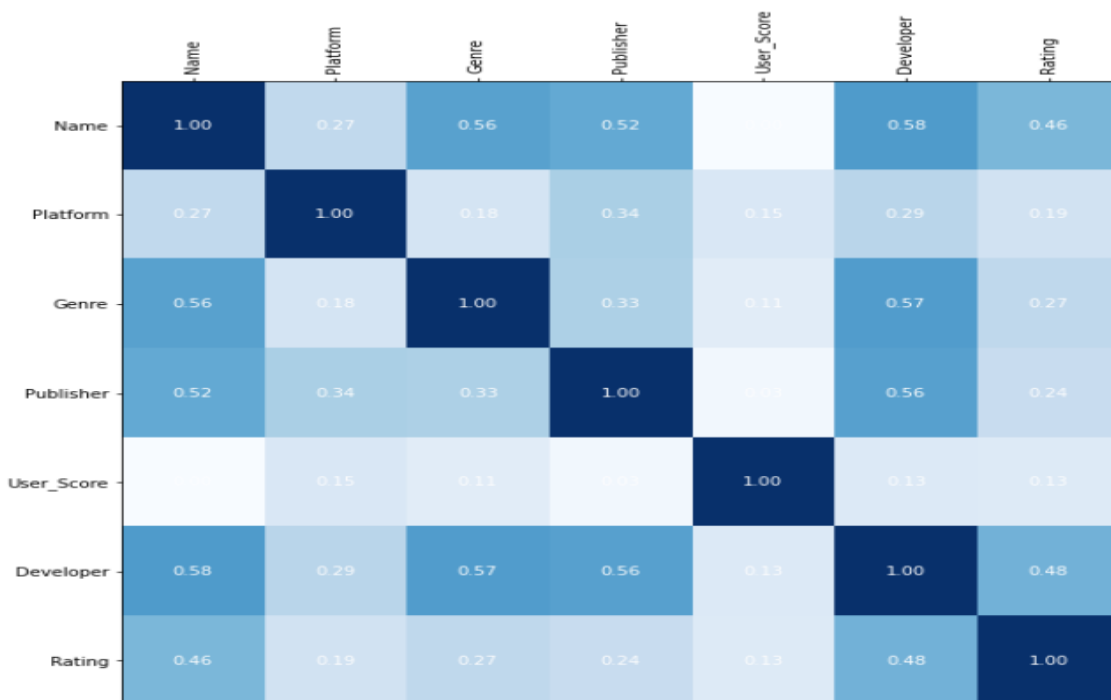
- Games Published Annually:
  - Peak at video games published in 2010, 2005. production then decreased until 2020 given this data, however after searching I found that the video games industry has grown exponentially in the past few decades.



- Correlation between numerical features

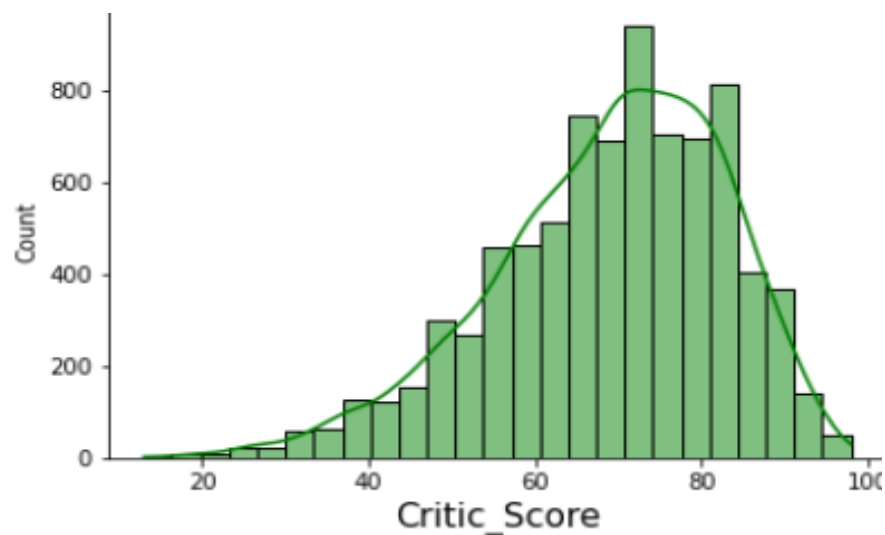


- Correlation between categorical features





- The features are not highly correlated with each other most of the matrix below 0.4
  - Sales columns only have a very high corr coefficient which indicates that high sales of a game in certain regions means high sales in others.
  - Correlation between critic count and critic score is not low but it's meaningless.
- Critic Score
  - A very important feature to predict sales of a game (more about this in the predictive question section).



## 4. Data Preparation

- Reading the data, there are several options that can be considered during this stage, such as:
  - Reading specific subsets of the data, rather than the entire dataset.
  - Selecting only columns with numerical or categorical data types, or both.
  - Applying encoding methods such as label, one-hot, or frequency encoding to transform categorical features into numerical representations that can be used in statistical analysis.
- Handle Missing values, there are several options that can be considered during this stage, such as:
  - Drop all missing values, this wasn't the best choice here.
  - We have a mix between categorical and numerical features:
    - For categorical, we always use the mode to fill the missing value.
    - For numerical, we had tested different approaches (mean/median/mode/KNN/Iterative imputer), thus the best one was the KNN imputation.
- Outliers detection:
  - We used the IQR to detect outliers in numerical columns, where the outliers will lie in the region below the lower bound ( $Q1 - 1.5 * IQR$ ) or above the upper bound ( $Q3 + 1.5 * IQR$ ).
- Outliers removal:
  - We considered:
    - removing outliers from a certain column.

- removing all outliers from all -specified- columns.
- removing common outliers from the specified columns. This was the technique used in the analysis.

## 5. Analysis

### 5.1. Descriptive Questions

Q1: For games rated more than X, what is the most popular genre ?

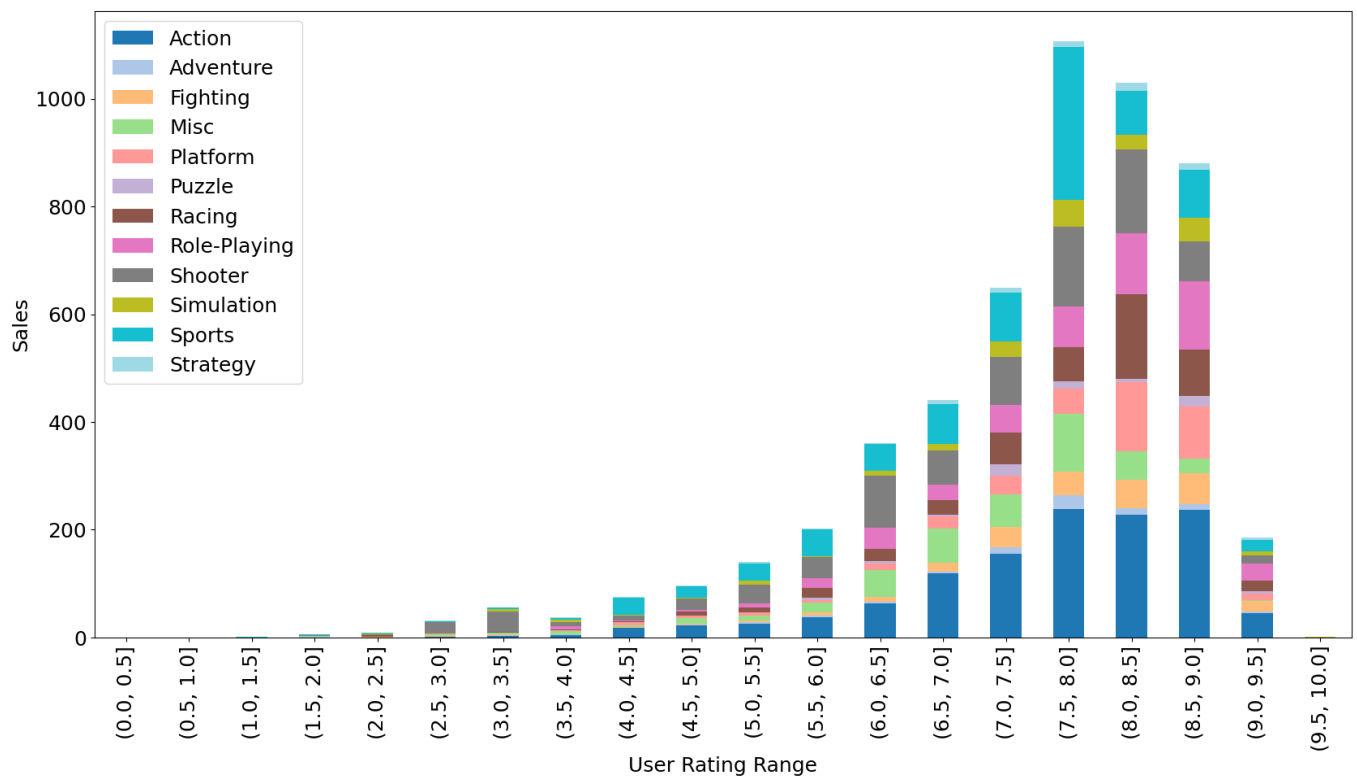
- Expectations

We expect that for each rating, there will be a genre that is most popular (has most sales for that rating)

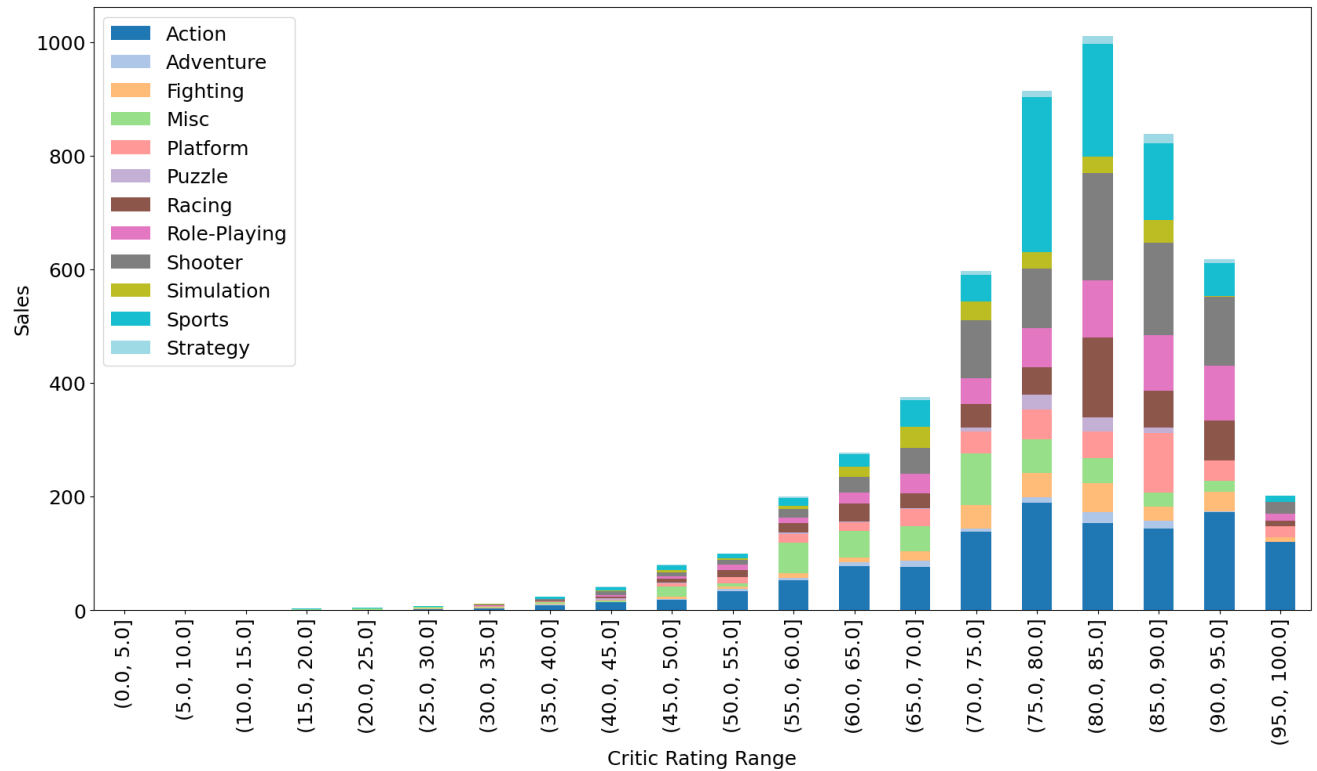
- EDA

We group ratings into bins of equal size, and for each bin, we find the genre that has highest sales

In the following graph: we plot user ratings vs sales for each genre



And here is the same graph but for critics score:



We notice that for ratings higher than  $\text{User\_Score}=8$  (and  $\text{Critics\_Score}=85$ ), the genre with most sales is *Action*. And for ratings more than  $\text{User\_Score}=7.5$  ( $\text{Critics\_Score}=70$ ), the genre with most sales is *Sports*. And so on.

## - Results and Insights

We thus conclude that for different ratings, there are different genres that have most sales.

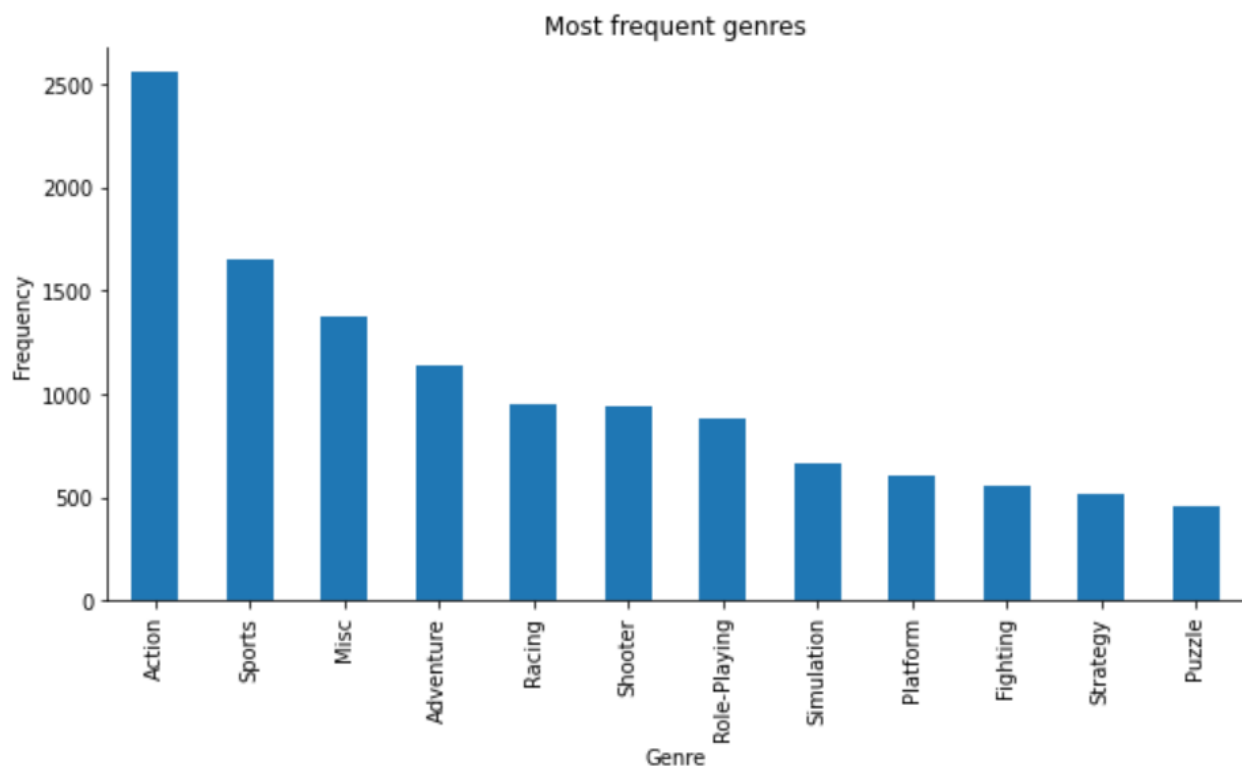
## Q2: Is there a certain genre that has the highest sales?

- Expectations

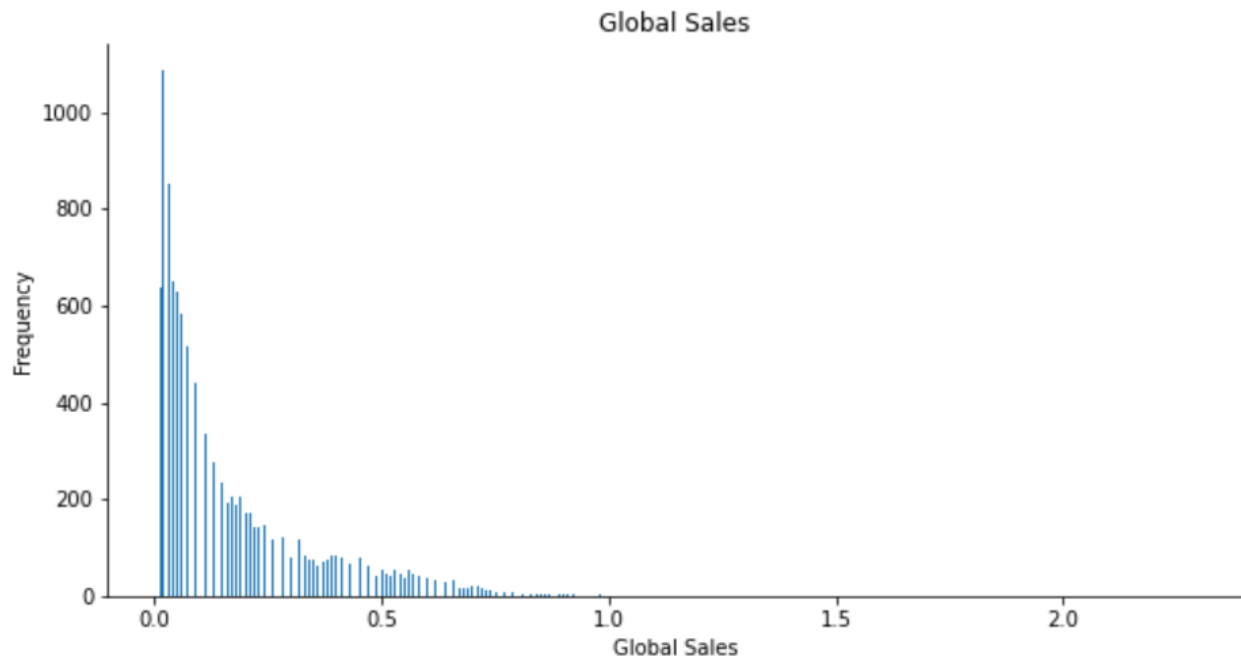
It's expected that in video games, There is a genre that is much preferred by the users than others and has the highest sales.

- EDA

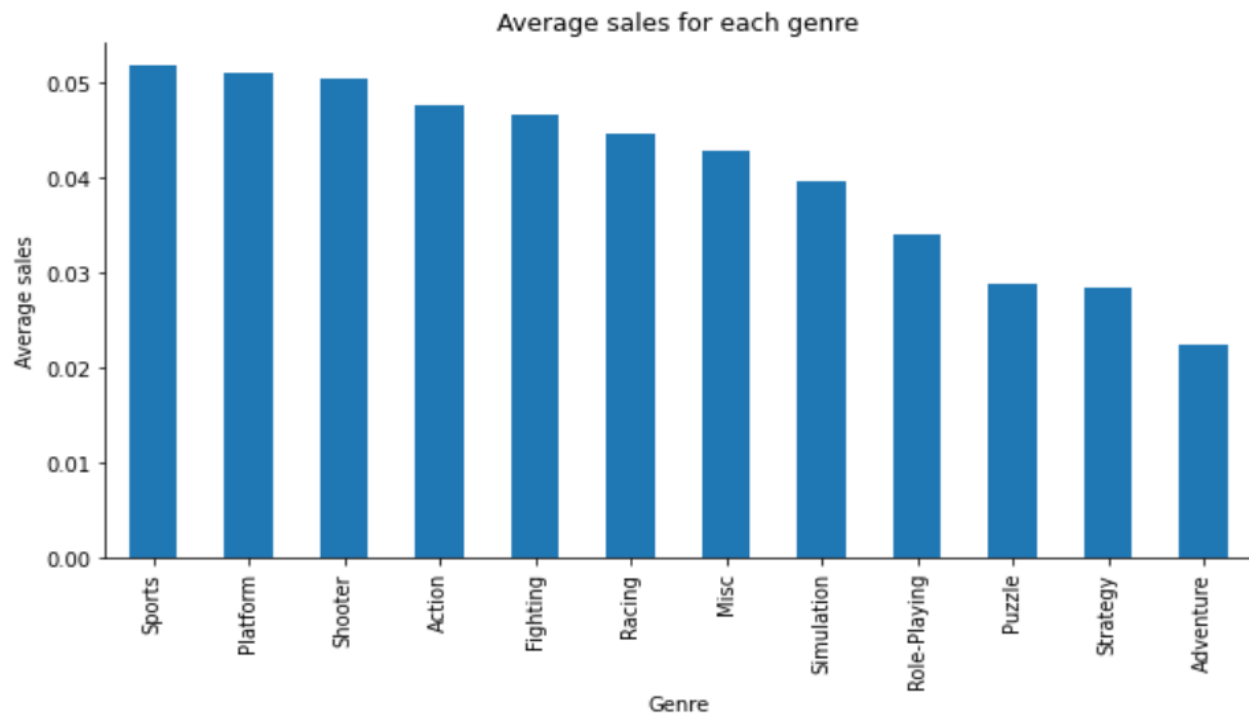
- Genre Distribution: Action genre is the top frequent genre in the dataset, followed by Sports and Misc genres.



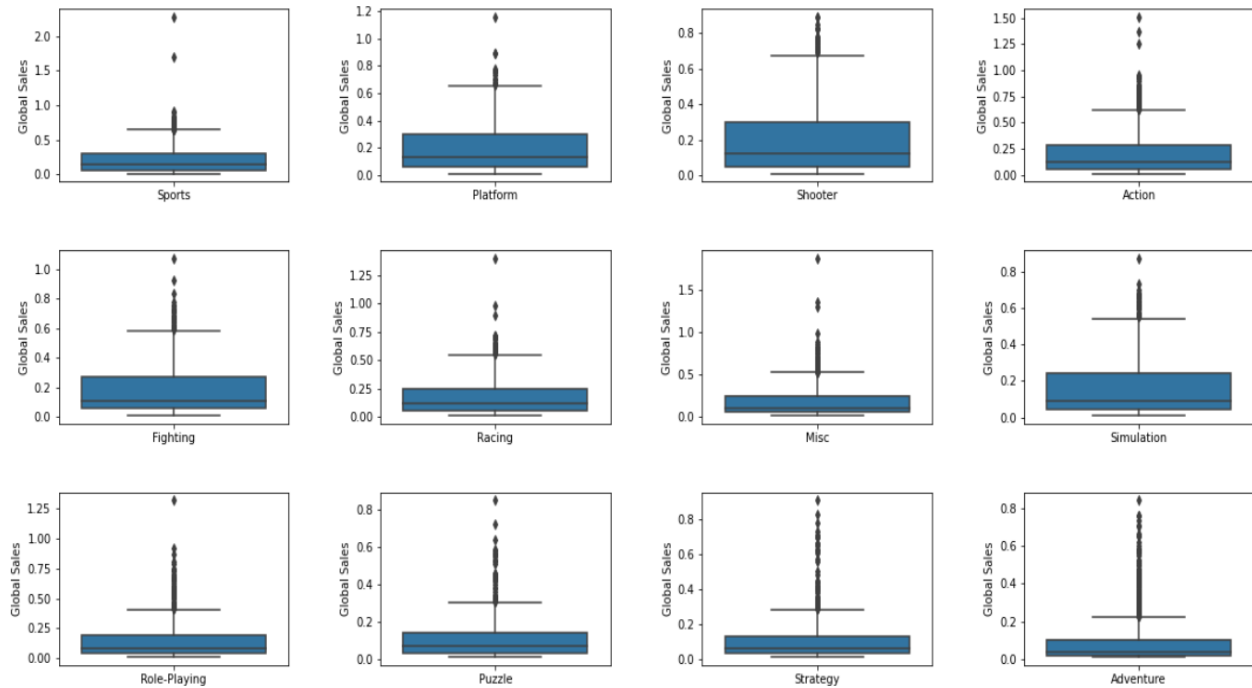
- Sales Distribution: The sales are in the range [0, 0.5] million dollars.



- Average sales for each genre: Genres from Sports to Fighting have almost the same average sales.



- Summary for average sales for each genre: Sports seem to have the highest values for outliers therefore it has the highest average sales.



	count	mean	std	min	25%	50%	75%	max
Genre								
Sports	1650.0	0.207497	0.190581	0.01	0.06	0.14	0.30	2.28
Platform	605.0	0.204512	0.188529	0.01	0.06	0.13	0.30	1.16
Shooter	942.0	0.202346	0.195428	0.01	0.05	0.12	0.30	0.89
Action	2555.0	0.191198	0.186796	0.01	0.05	0.12	0.28	1.51
Fighting	552.0	0.186413	0.186754	0.01	0.06	0.11	0.27	1.07
Racing	954.0	0.178721	0.168939	0.01	0.05	0.12	0.25	1.40
Misc	1373.0	0.172017	0.180285	0.01	0.05	0.10	0.24	1.87
Simulation	660.0	0.159000	0.163422	0.01	0.04	0.09	0.24	0.87
Role-Playing	879.0	0.136246	0.153386	0.01	0.04	0.08	0.19	1.32
Puzzle	453.0	0.116645	0.130358	0.01	0.03	0.07	0.14	0.85
Strategy	512.0	0.113926	0.143068	0.01	0.03	0.06	0.13	0.91
Adventure	1139.0	0.089860	0.118589	0.01	0.02	0.04	0.10	0.84



- Results and Insights
  - Expectation matches the data.
  - Sports genre has the highest sales.
  - The sales are in the range  $[0, 0.5]$  million dollars.

## 5.2. Exploratory Questions

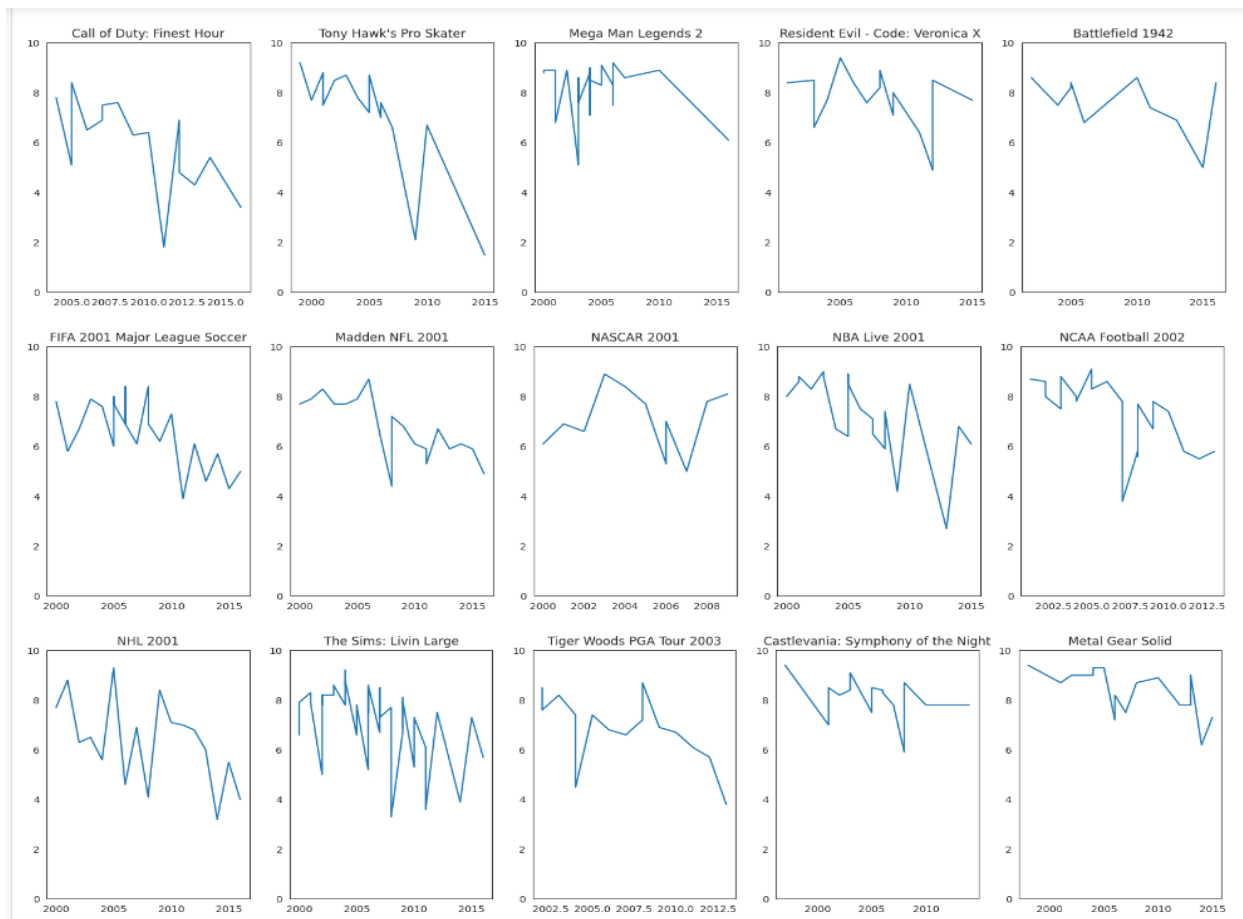
**Q1: For some specific game of multiple versions, Does rating get better or worse over time?**

- Expectations

We expect that ratings increase as developers notice bugs and fix them. Also they add new features and enhancements.

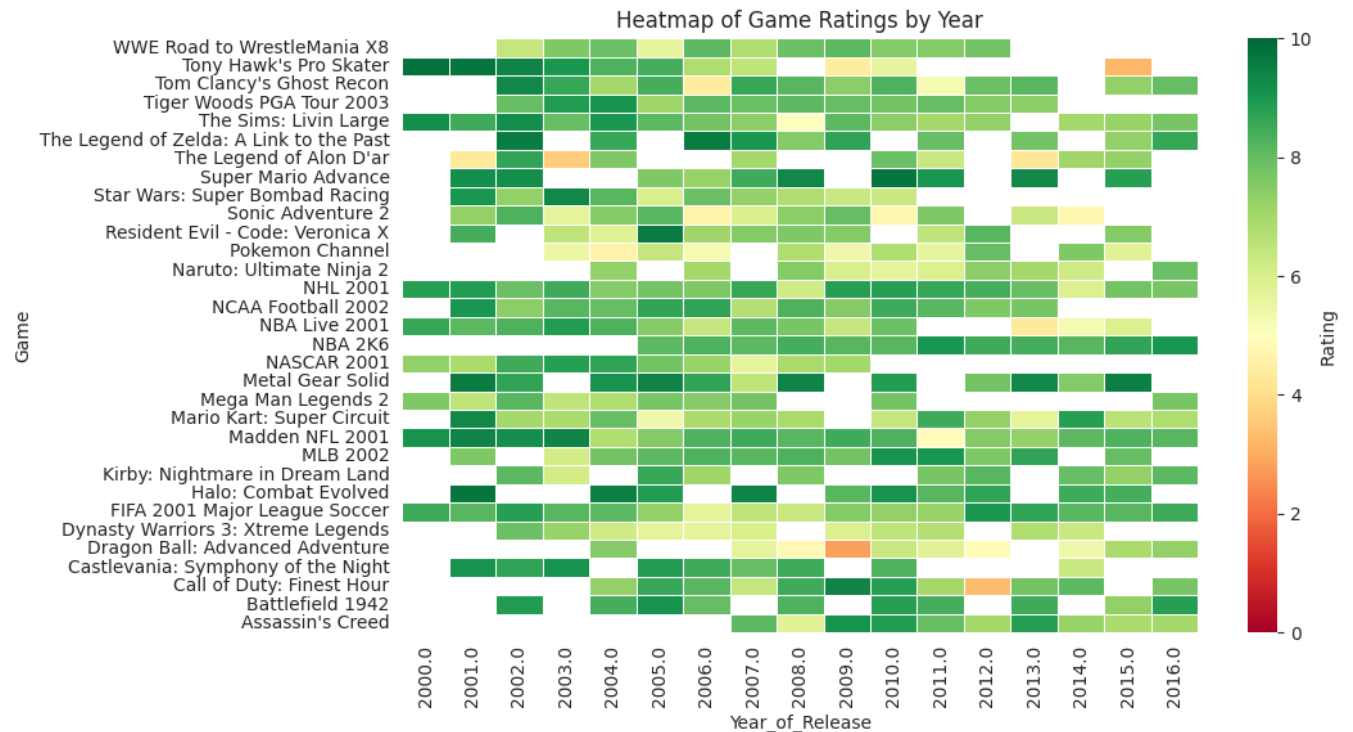
- EDA

We group games by publisher and consider games that have the same first syllable as the same game but with multiple versions. We then plot these games' ratings over the years.



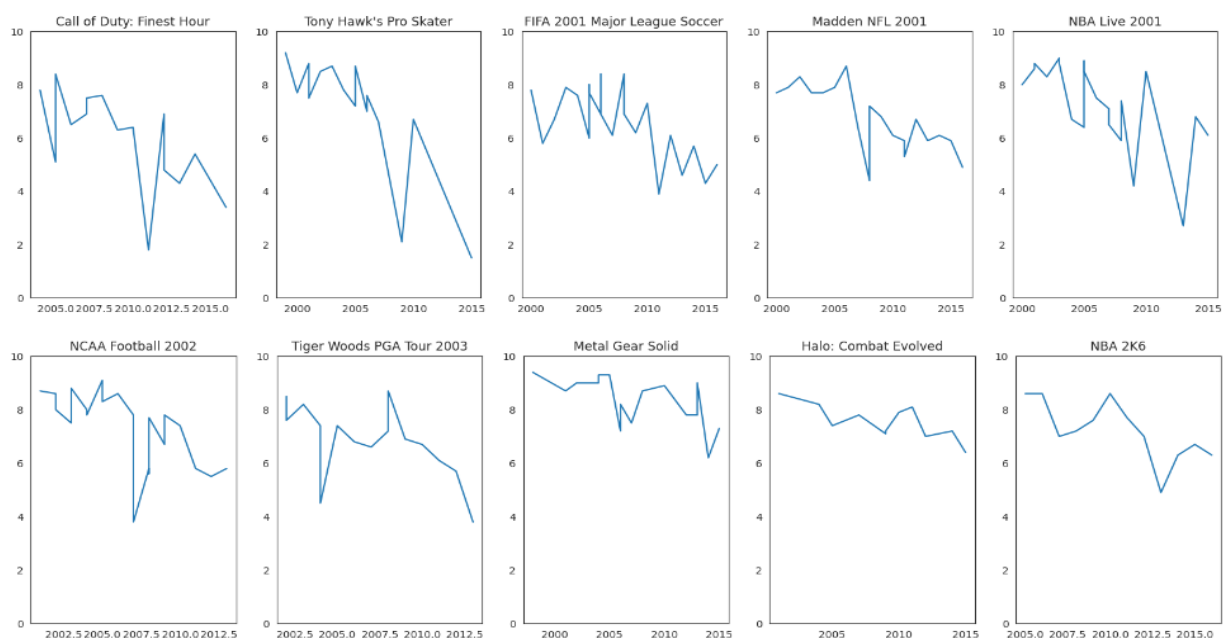
The previous graph shows a subset of games, where each game has its rating drawn vs year. Games chosen in it are those that have more than 10 versions (10 ratings over 10 years)

We notice that for most games, their average ratings tend to be lower over time. (fluctuating but towards decrease)

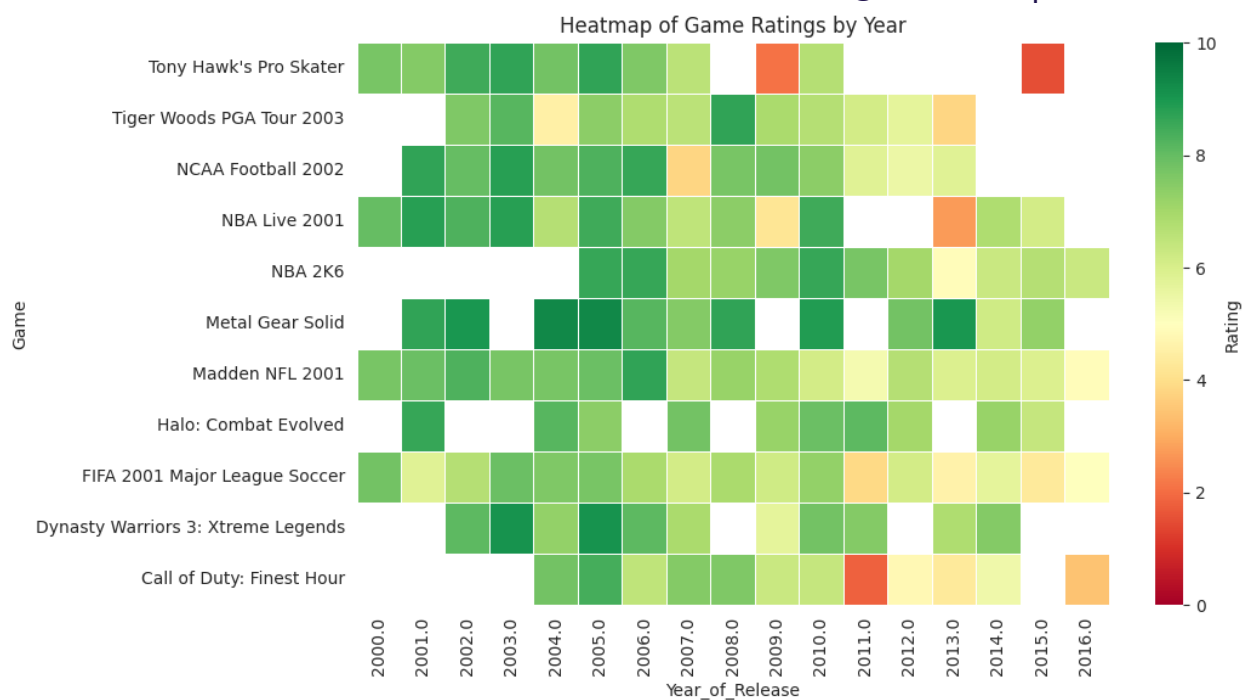


This heatmap has 32 games. We notice that on the right, color is slightly lighter, which means that ratings slightly decrease over time.

The following games are drawn with their heatmap. They are games that their ratings have correlation with years (found by hypothesis testing) - their percentage is 43% of games that have more than 10 games



We can also visualize this in the following heatmap:



The graphs for critics scores are similar.

- Hypothesis Test

$H_0$ : there is no correlation between year of release and average user score for all games

$H_A$ : there is a correlation between year of release and average user score (positive or negative)

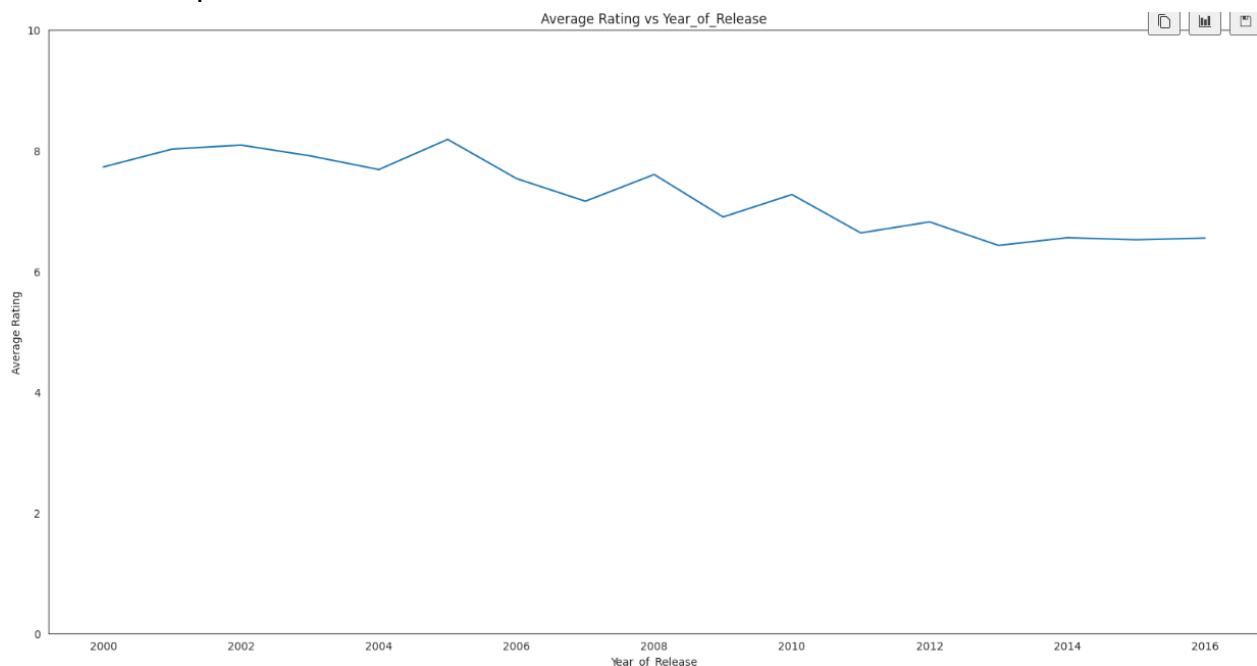
We find average rate for each year for all games that have a version in that year. Then we compute pearson correlation coefficient between all years and average rating. We use the p-value of the test of correlation  
Set  $\alpha = 0.05$

$pvalue = 3.75 \times 10^{-7}$ ,  $corr = -0.91$

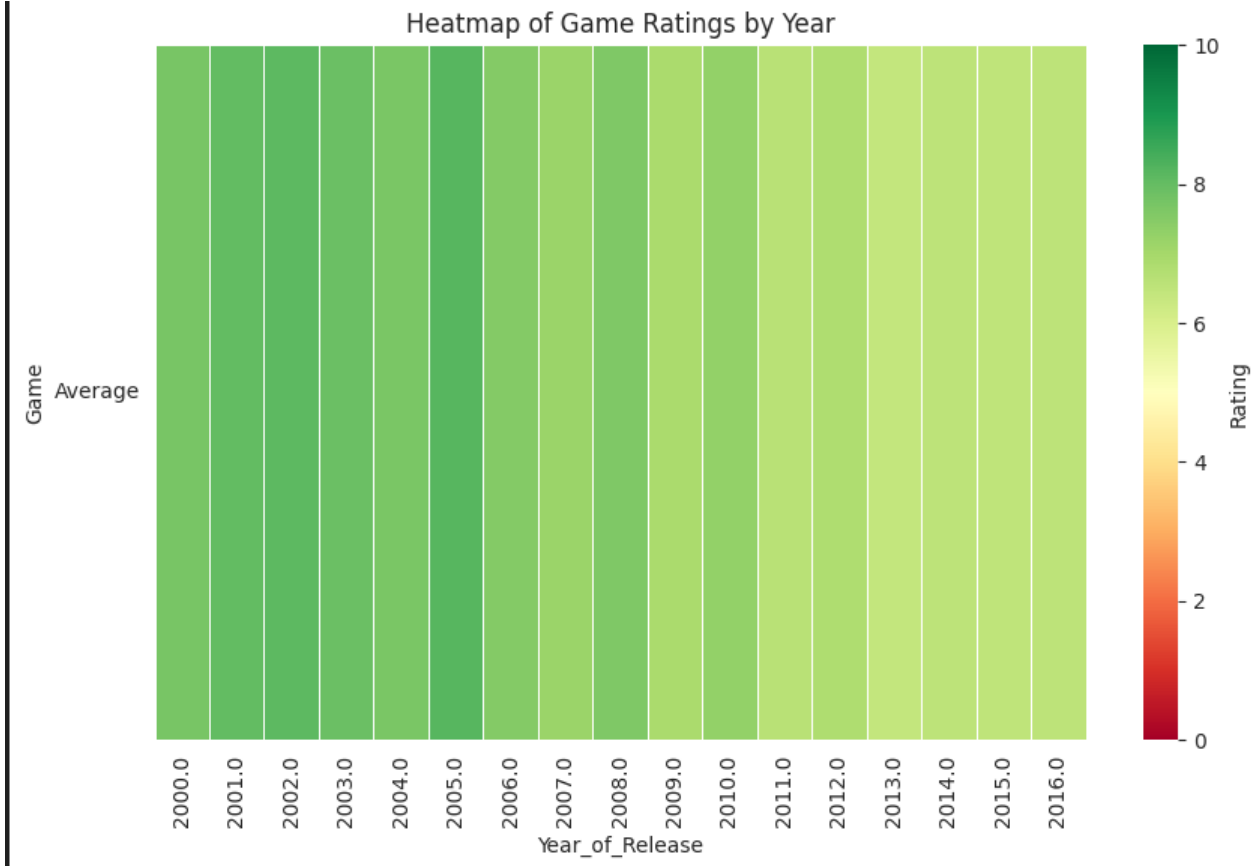
Thus we conclude that on average, ratings DECREASE over years

- Results

When we plot average ratings vs years:



We see a slight decrease in average ratings over the years.  
And here is the corresponding heatmap



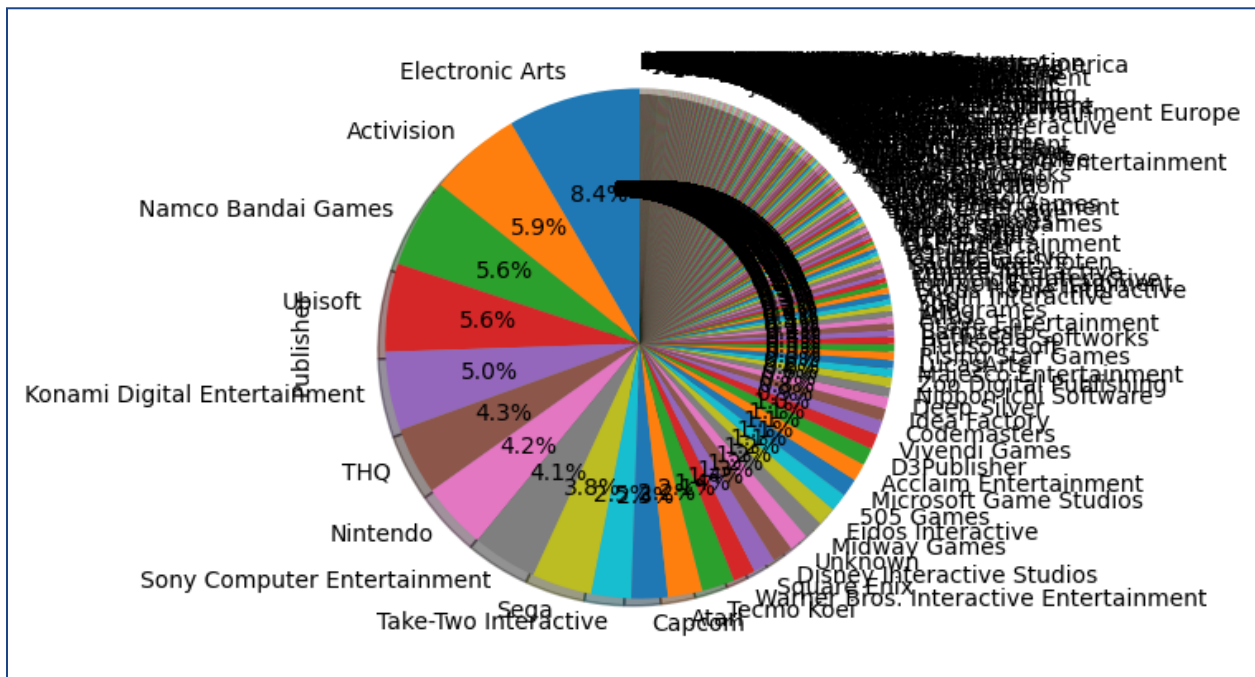
Q2: Does the name of the publisher have a role in increasing the sales?

- Expectations

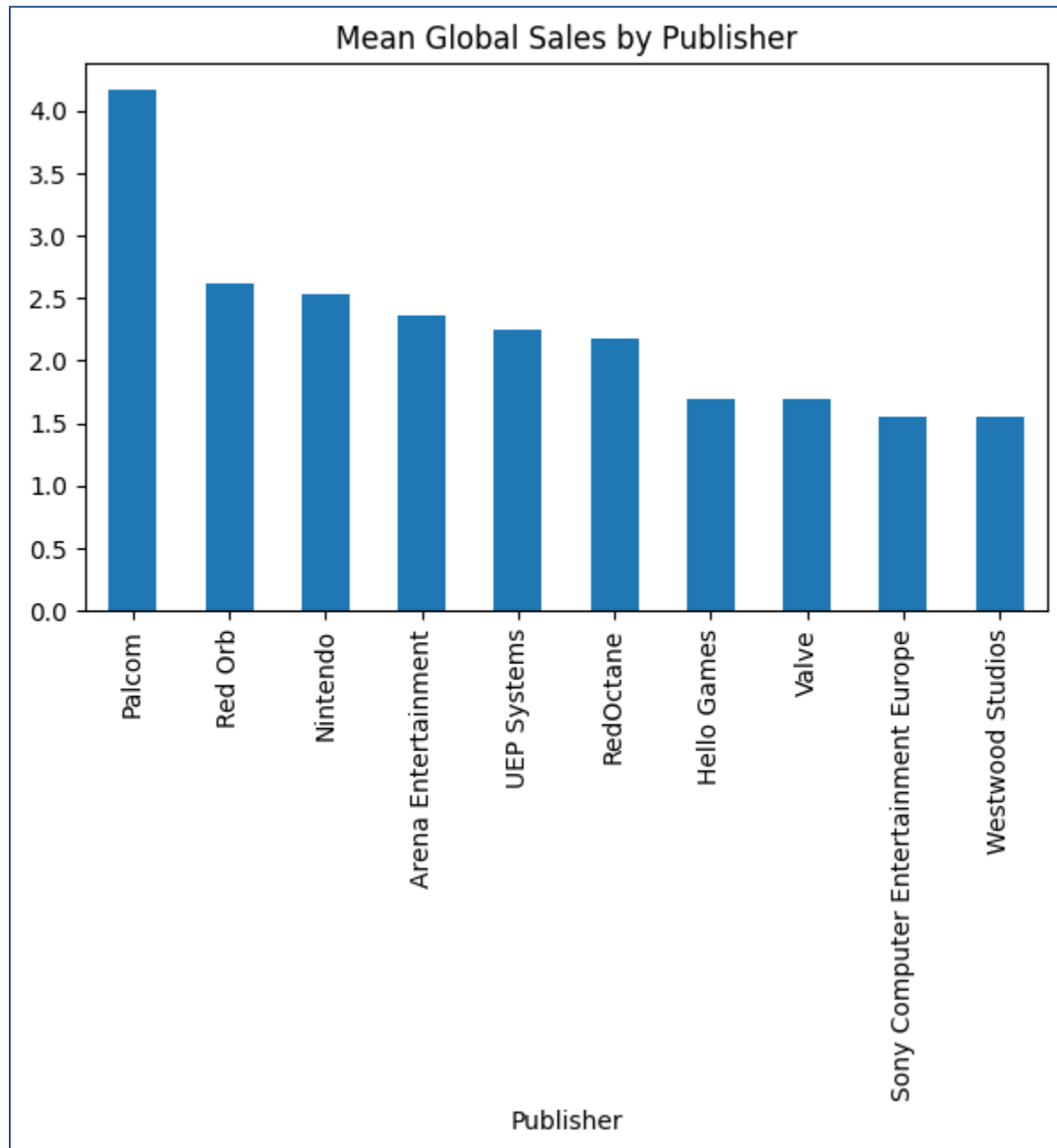
Yes, publisher name can affect video game sales. According to a study by CAUSEweb, publisher name is one of the factors that impact video game sales<sup>1</sup>. Publishers help developers navigate the business of game development by offering various services such as public relations and marketing, negotiating with platform holders, and administrative work

- EDA

- Group Global sales by publisher name: we have too many publishers!! many of these publisher has only one game published in our data so it's not correct to consider them in this analysis, specially that one of them with the highest sales. Drop publisher with very small number of games to avoid misleading

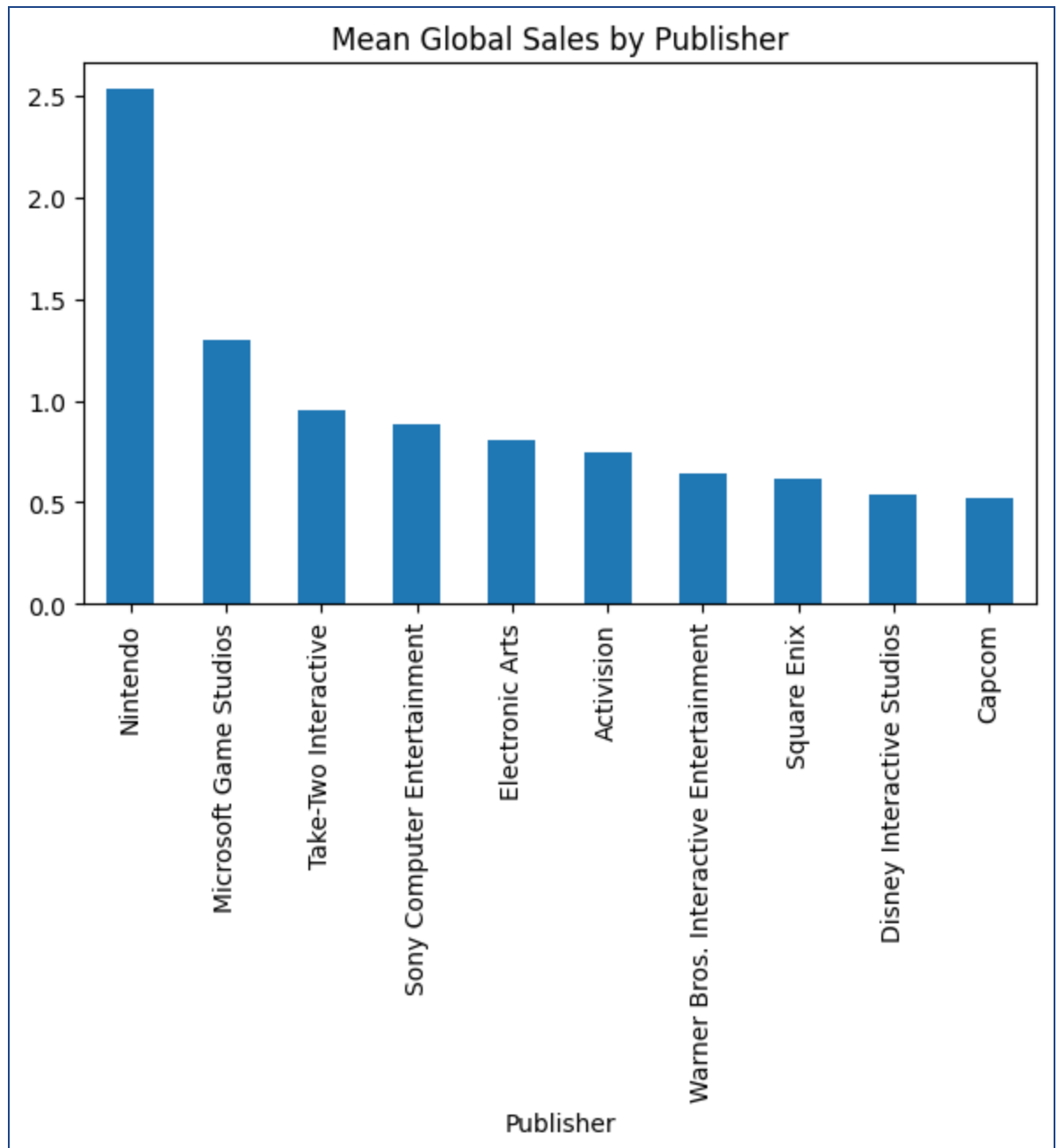


-Before Dropping:



- After Dropping: Result is significantly different so we avoid incorrect results





- Hypothesis Test
  - H0: Publisher has no effect on the global sales of a game.

-H<sub>A</sub>: Publisher affects the global sales of

- Using ANOVA Test: P-value =  $2.283678202155225e-10 < 0.05$  so reject null hypothesis

- Results and Insights

- Publisher of a game has an effect on its global sales.

### Q3: Compare platforms based on how long they stay competitive in the market.

- Expectations

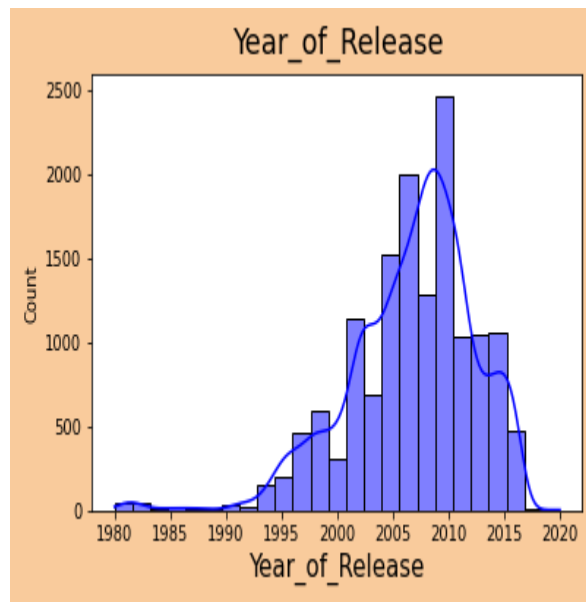
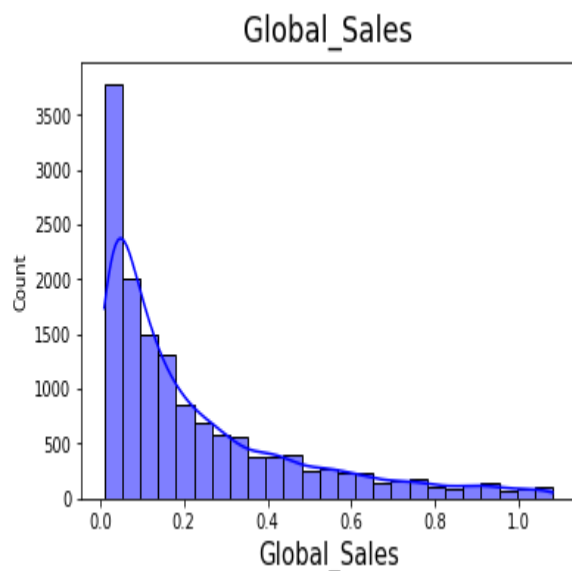
New platforms get the hype and keep increasing until they peak and then they start decreasing and by the time they are old, newer platforms take their share and repeat the cycle.

- EDA

Relevant columns: ['Platform', 'Global\_Sales', 'Year\_of\_Release']

Nulls handling: Dropped as they are a very small fraction (269 / 16719)

Outliers: the column 'Global\_Sales' has mean of 0.53 but magically also has values more than 20 !!, so we filter about 1872 rows and now have a more reasonable mean of 0.22 and max of 1.08

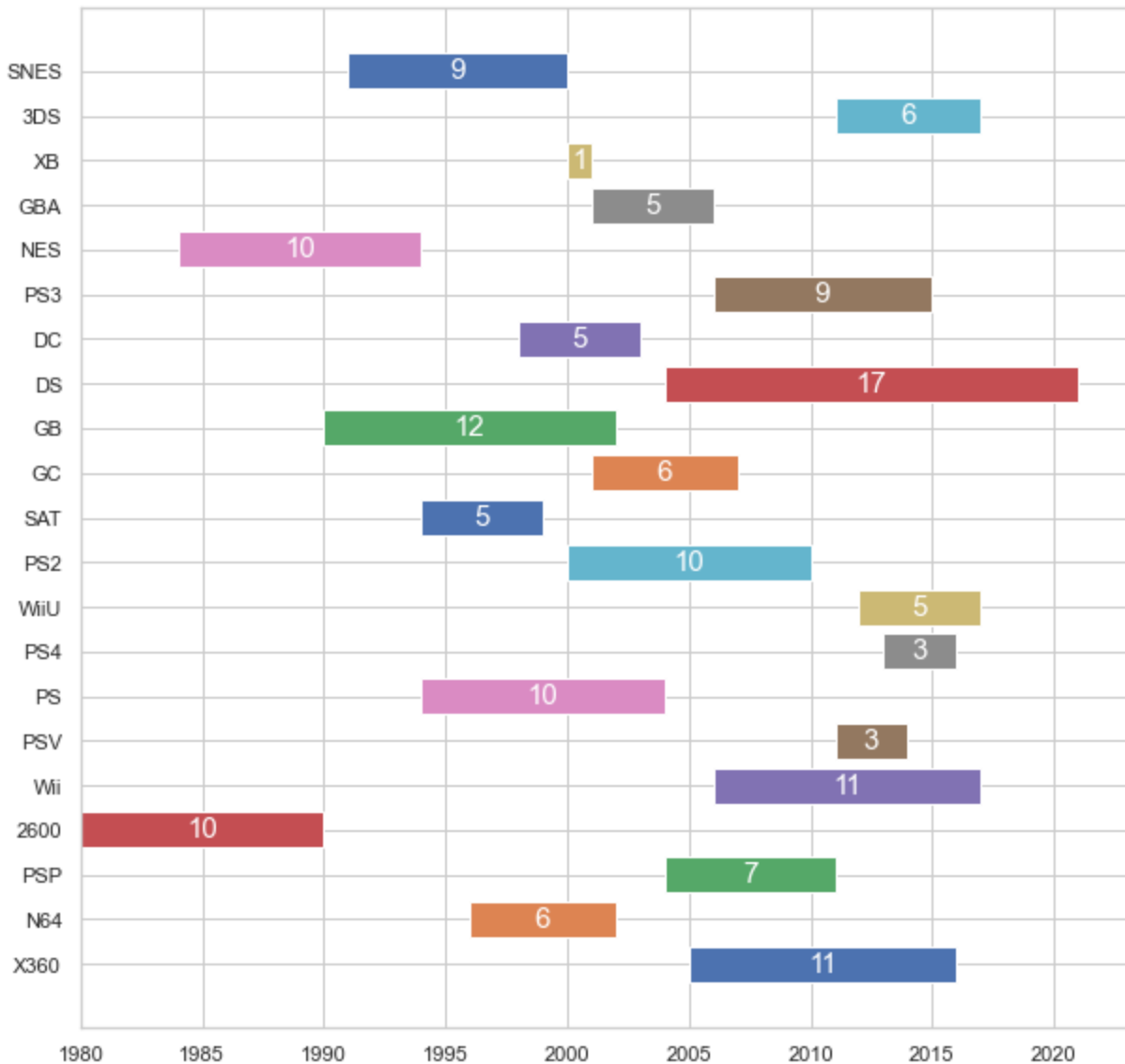


Next step is that we group all rows, to convert it from representing a game to representing Global\_Sales of specific platform in a specific year

Platform	Year_of_Release	Global_Sales
WS	2000.0	0.340000
PS2	2004.0	0.296866
PC	1997.0	0.010000
WiiU	2013.0	0.260278
SNES	1996.0	0.287692
PS4	2015.0	0.239189
PC	1999.0	0.248000
PC	2004.0	0.062500
PC	2001.0	0.117692
PS3	2009.0	0.351716

Now we do filtering once more, platforms with too few num\_years are misleading and hard to explore, so we drop all platforms with no sales in less than 5 years, which ends up dropping 10 platforms.

Now we Need to extract the active period for every platform we can say that active is from first year it started achieving an average of at least half of its peak\_mean to the last year it did, this threshold could be improved by utilizing standard deviation and it is considered in future improvements.



We can see that the maximum difference is between DS and XB which is a whopping 16 years

- **Insights**

We extracted how many active years each platform has achieved, however our dataset is limited by the time slot it captured, i.e.: some platforms were active from even before the very first game in this dataset, some platforms kept active after and of course some platforms didn't have enough games to showcase every single year of every single platform.

Q4: Do users or critics rate a specific platform or genre higher than others? is actually 4 questions disguised as one

- Expectations

No, users or critics should not be biased towards a specific platform or genre.

- EDA

Relevant columns: User/Critic score, User/Critic count and Platform/Genre

Invalid Values: User\_Score was considered an 'object' and it had 'nan' entries, so it was converted to 'float64' and 'nans' converted to zeros.

Nulls handling: Some games had a score of 'zero', this basically meant it wasn't rated, However, we didn't filter them as we will filter upon User/Critic count and they will disappear.

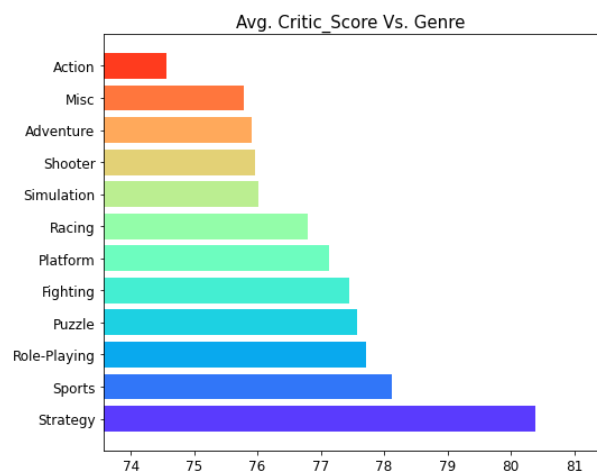
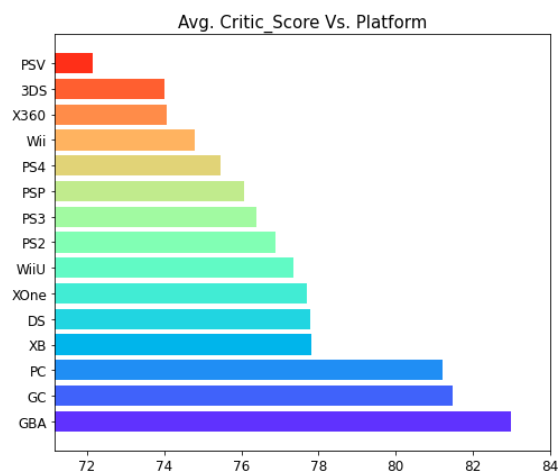
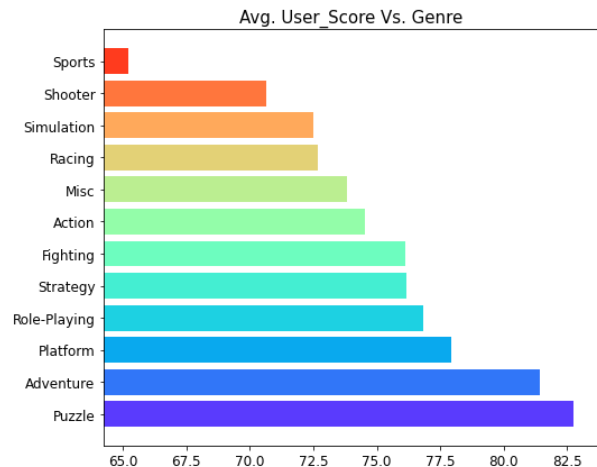
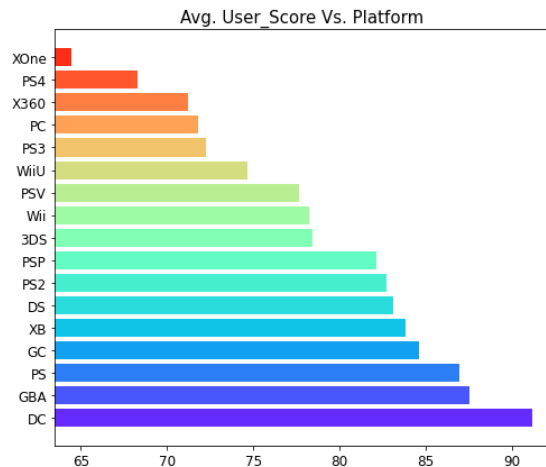
Platform and Genre both have no nulls or invalid values.

Filtering on User/Critic Count:

User_Count							
-----							
>= 0	>= 10	>= 30	>= 60	>= 70	>= 100	>= 500	
-----+-----+-----+-----+-----+-----+-----+							
7017	5342	3278	2218	2038	1629	537	
Critic_Count							
-----							
>= 0	>= 10	>= 30	>= 35	>= 40	>= 50	>= 70	
-----+-----+-----+-----+-----+-----+-----+							
7017	5891	2665	2097	1672	987	285	

We chose a threshold of 70 for User\_Count and 35 for Critic\_Count, So, we stayed with around 2K rows to continue analysis with stronger confidence of scores.

- Results and Insights



Clearly, Bias DOES exist, however, and we used t-test to test its significance

## • Hypothesis Test

Use T-test to test the relationship between a numerical variable (User/Critic Score) and a categorical variable (Platform / Genre) we want to test if there really exists bias in scoring.

- null hypothesis: Bias doesn't exist, Platform/Genre doesn't affect the score
- alternative hypothesis: Bias exists, Platform/Genre does affect the score

```
For User_Score Vs. Platform:
worst value when comparing PS2 Vs. XOne
t-statistic  = 12.448478
Reject H0
-----
For User_Score Vs. Genre:
worst value when comparing Sports Vs. Role-Playing
t-statistic  = -9.198338
Reject H0
-----
For Critic_Score Vs. Platform:
worst value when comparing X360 Vs. PC
t-statistic  = -7.510106
Reject H0
-----
For Critic_Score Vs. Genre:
worst value when comparing Action Vs. Strategy
t-statistic  = -4.274801
Reject H0
-----
```

- Interpretation of Results

We can say with absolute confidence that  $H_0$  was wrong and that Bias does exist. That is, \*Users\* and \*Critics\* are biased towards specific \*Platforms\* and \*Genres\*.



## Q5: Which regions prefer which genres?

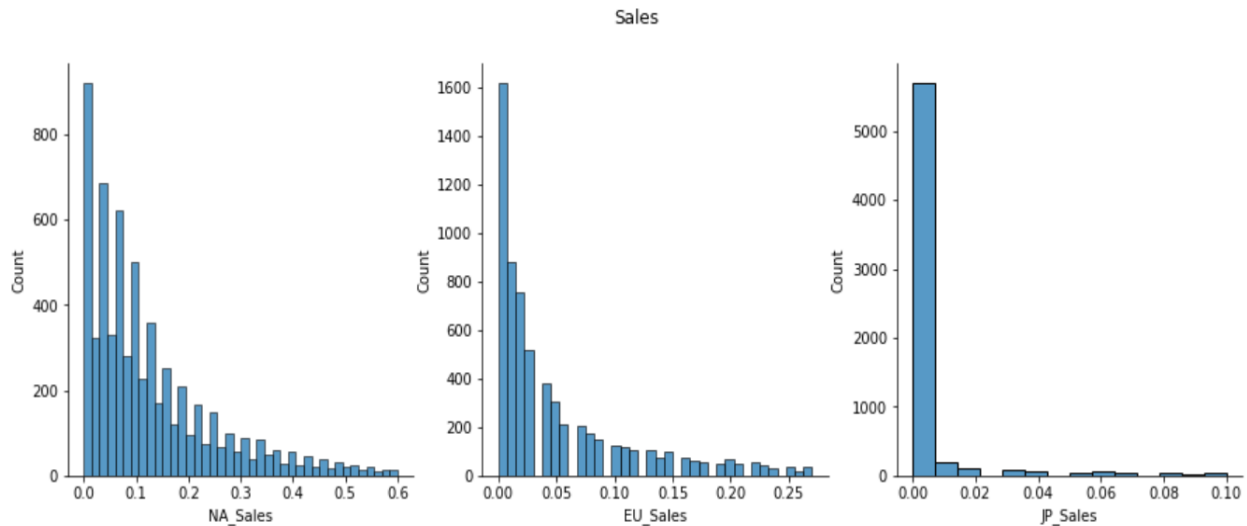
- Expectations

It's expected that each region prefers a certain genre as all different cultures (UA, NA, JP) don't seem to prefer only one genre.

- EDA

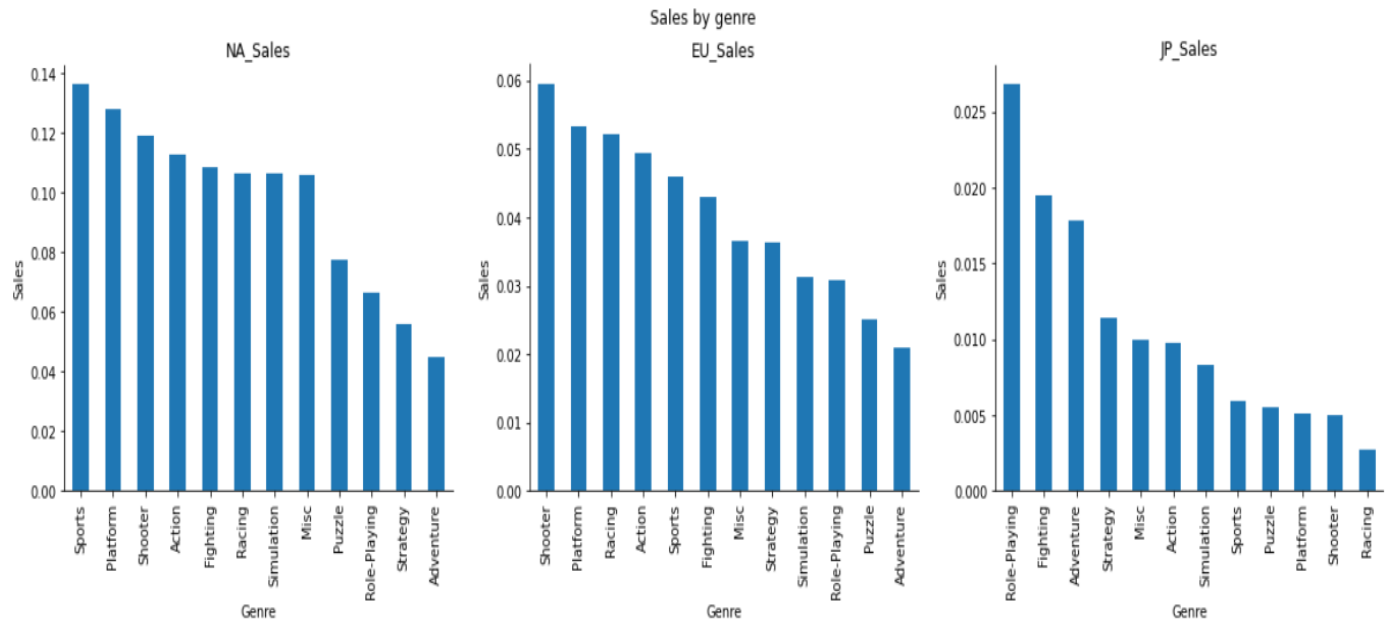
- Each Region sales' distribution:

- We can see that North America has more sales than Europe and Japan.
- Japan has the least sales.

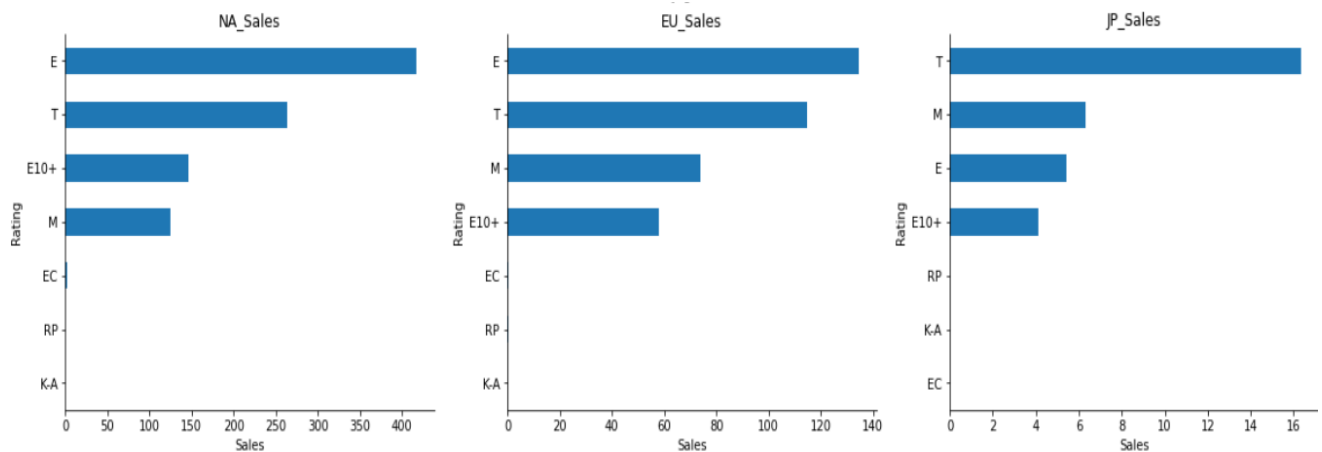


- Sales by genre in each region:

- For NA, the highest sales are for the Sports genre.
- For EU, the highest sales are for the Shooter genre.
- For JP, the highest sales are for the Role-Playing genre.



- Impact of ESRB rating on the sales: We'll study whether the ESRB can have an indirect effect on the sales for each genre.
  - for North America, the highest sales are from those games which are available to everyone.
  - for Europe, the highest sales are from those games which are available to everyone then teenagers.
  - For Japan, the highest sales are from those games which are rated for teenagers.



- Most common genres with ratings E,T,M:
  - The most popular genre with rating E is Sports then Racing.
  - The most popular genre with rating T is Action then Role-Playing and shooter.
  - The most popular genre with rating M is Action then Shooter.



## ● Hypothesis Test

To test the relationship between a numerical variable (NA/EU/JP Sales) and a categorical variable (Genre) we want to test if sports is the preferred genre in all regions.

Let  $\alpha=0.05$

- $H_0$ : All Regions prefer the same genre (genre with highest global sales -Sports-).
- $H_A$ : Each Region prefer different genre

- T-Test:
  - For NA, Retain H0.
  - For EU & JP, Reject H0.
- Anova Test:
  - For NA, Retain H0.
  - For EU & JP, Reject H0.
- Results and Insights
  - Expectation matches the data.
  - For North America, the most popular genre is Sports.
  - For Europe, the most popular genre is Shooter.
  - For Japan, the most popular genre is Role-Playing.
  - For the impact of ESRB ratings:
    - The ESRB rating may play a role in influencing the sales of video games.
    - In North America, the Sports genre with an E rating dominates the market and generates the highest sales.
    - In Europe, games rated for Everyone and Teenagers are the most popular, with the Shooter genre being the most common with a T rating, and generating the highest sales. This trend in Europe suggests a preference for games with violent elements that may not be suitable for children.
    - In Japan, games rated for Teenagers are the most popular, with the Role-Playing genre being the most common with a T rating and generating the highest sales.

## Q6: Is there a certain publisher whose sales are most coming from a certain region?

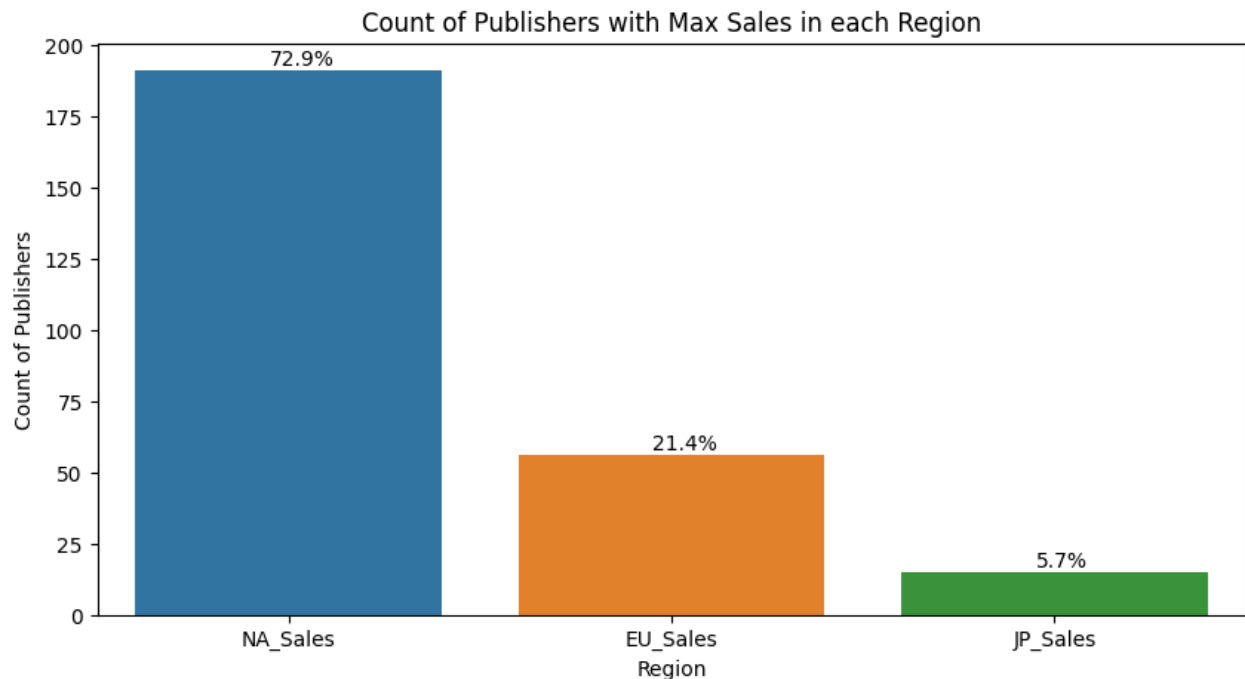
- Expectations

For most of the publishers, the sales are coming from some specific regions.

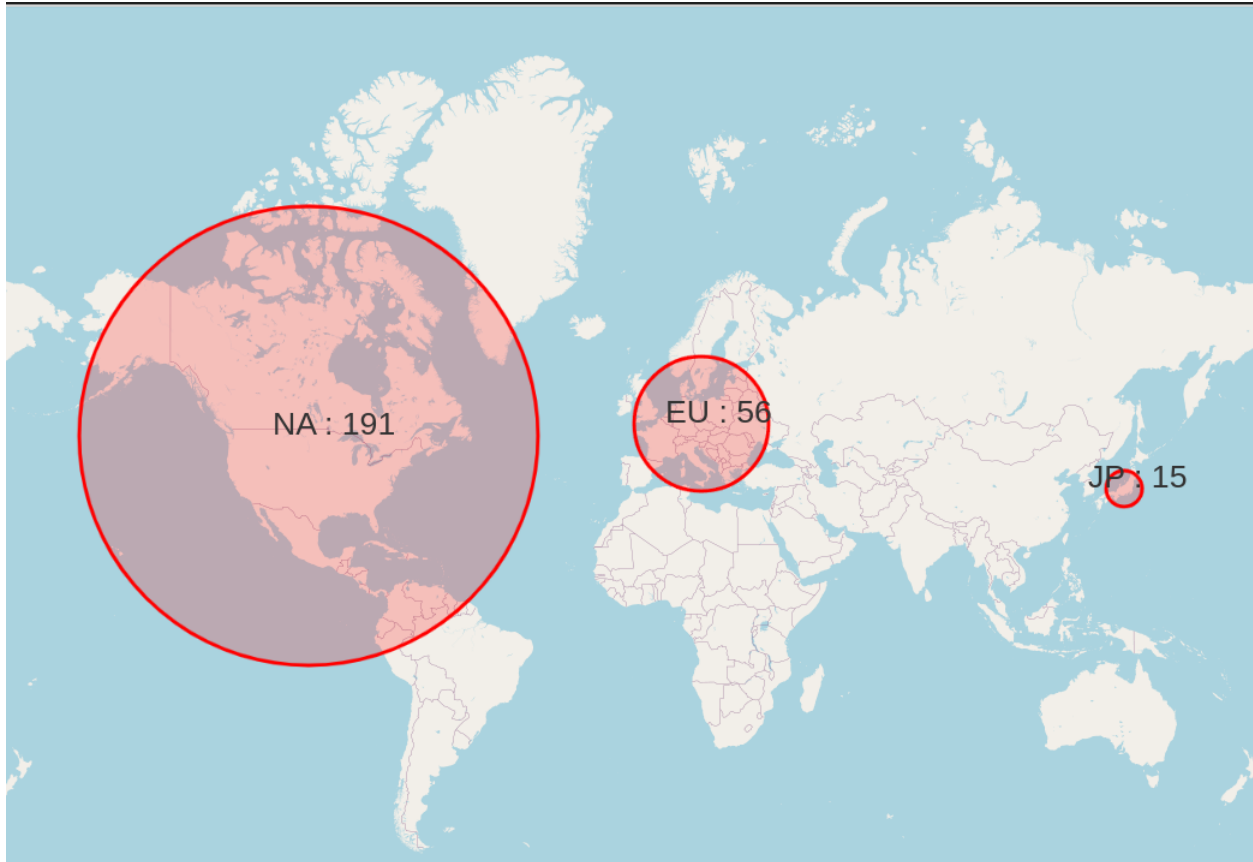
- EDA

We group the dataset by Publisher, then sum sales of each region for this publisher. We find region that has maximum sales for each publisher.

This histogram shows the percentage of publishers that have their max sales in each region.



And here is a map that visualizes their counts



- Results and Insights

We notice that North America gives maximum sales for 72% of the publishers (191 publishers). Followed by Europe: 21.4% of the publishers have their max sales coming from it. Then Japan (only 5.7%)

So, we conclude that it is advisable to invest in North America and increase marketing for games there.

## 5.3. Predictive Questions

Q1: Predict sales of a game given release date, platform, publisher and developer.

- Expectations

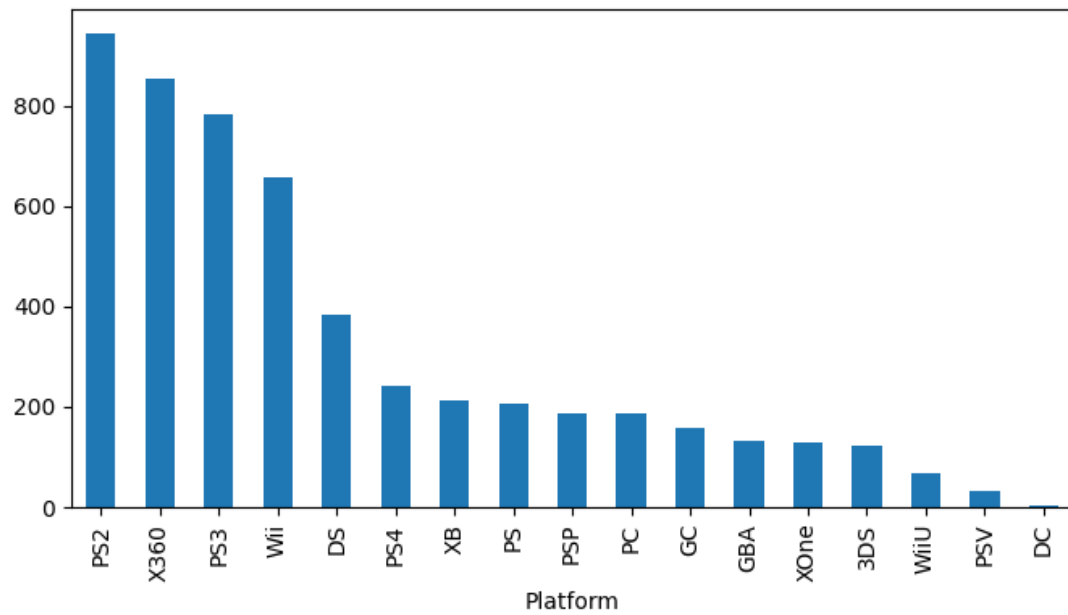
The model should determine the game characteristics through year, developed, publisher and platform that best predicts the sales of a game.

- EDA

- Global sales distribution over all the genres

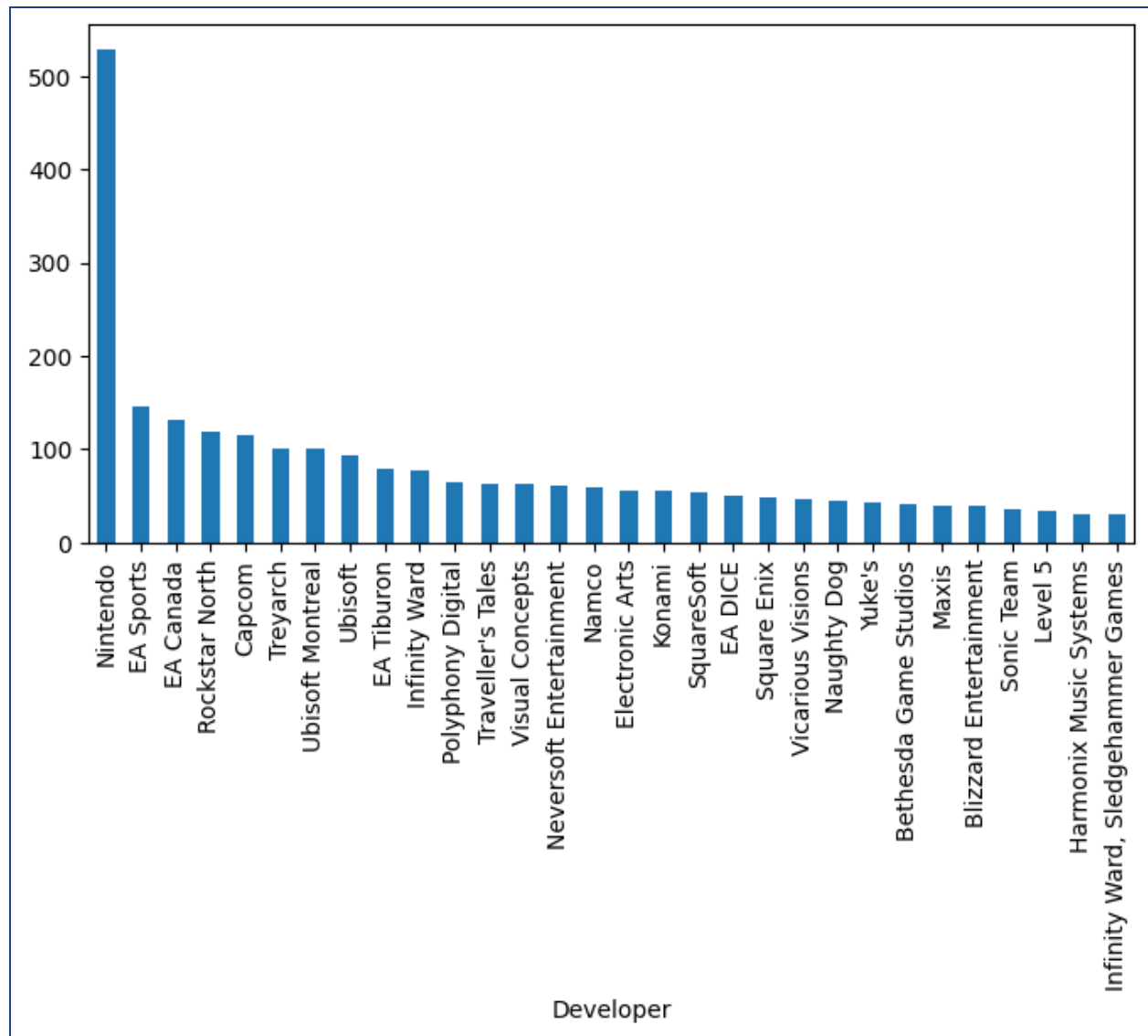


- Platforms with highest global sales

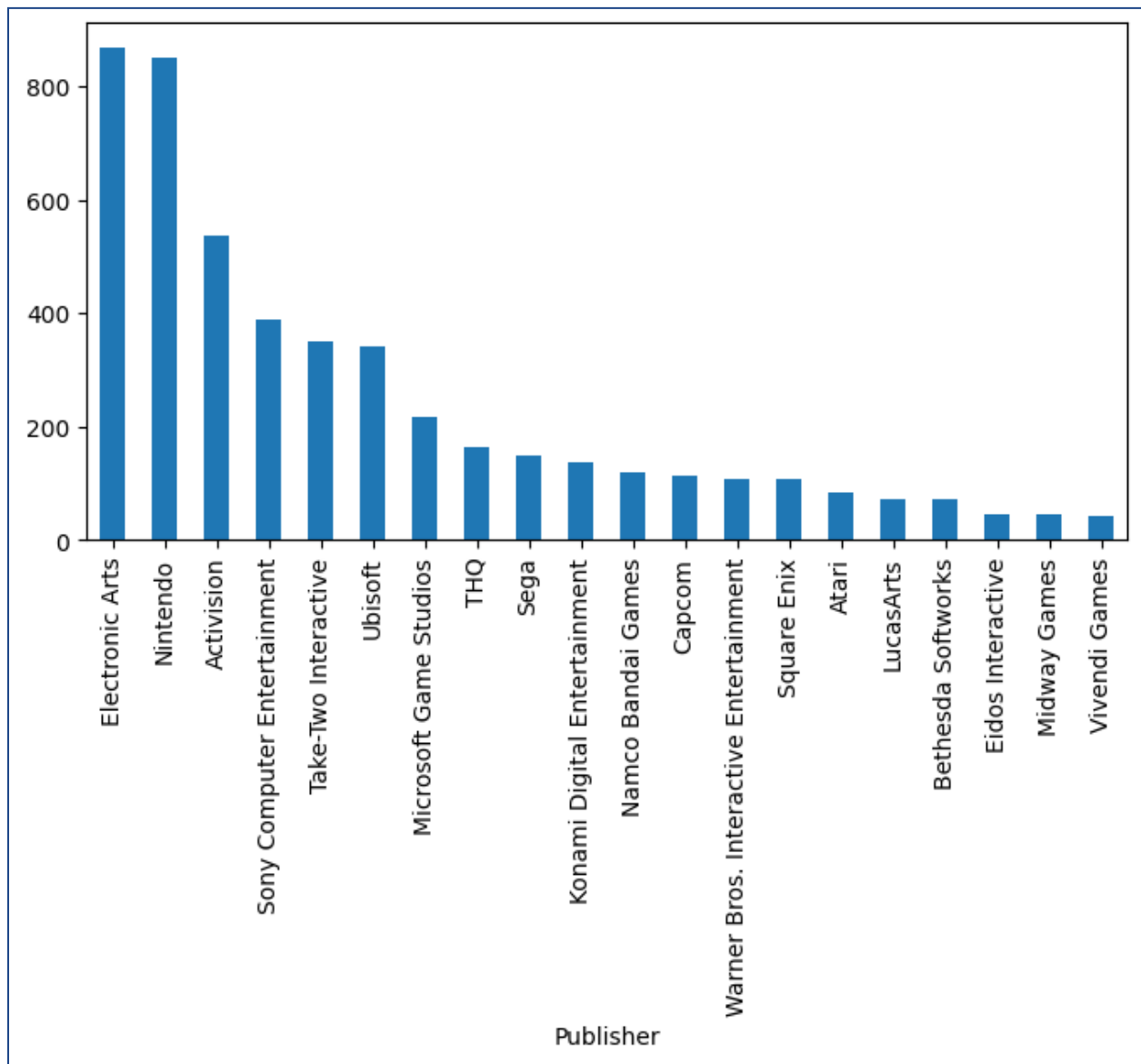


- Top 30 developers with highest global sales



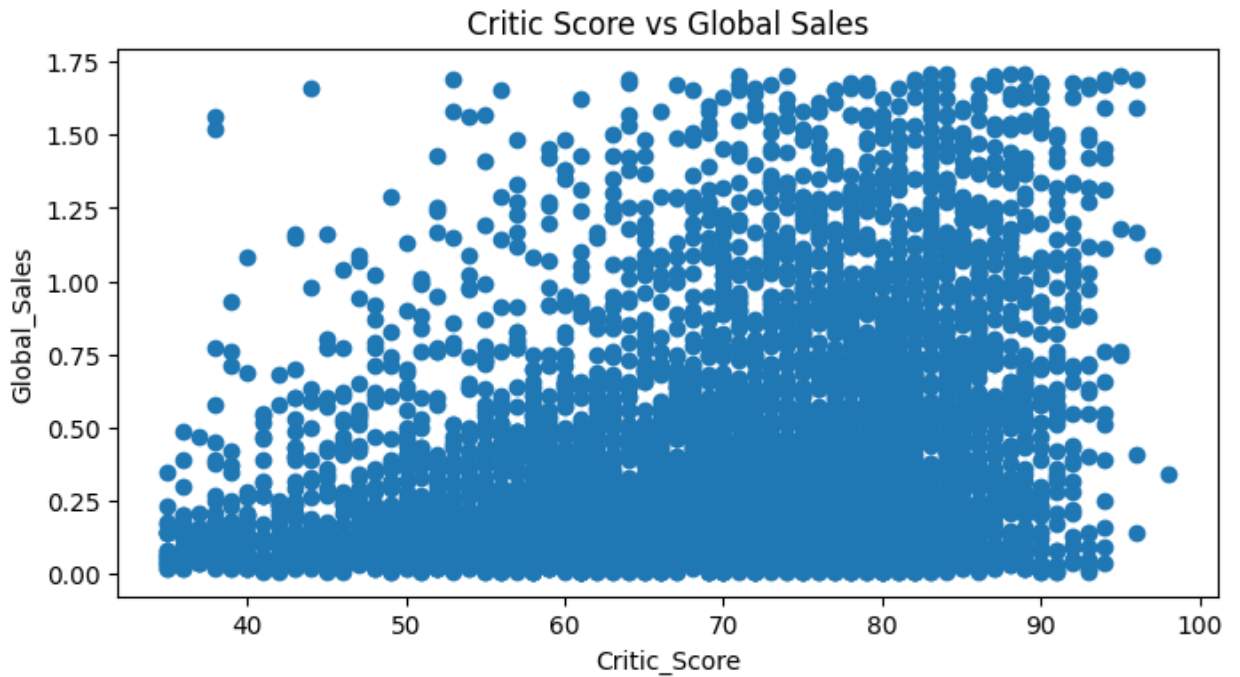


- Top 20 Publishers with highest Sales



☑ **Results:** There is significant difference between mean sale for different categories features: platform, features, developer so they may affect sales and help model predict sales, will see!

- **Global sales VS Critic Score:** it seems that there's a correlation between critic score and sales. So this feature will help the model predict the sales



- **Regression Models:**

Model	RMSE	R2-Score
Catboost	0.227	0.63
Catboost (with keeping outliers)	0.7	0.87
SVR	0.38	0.12
Elastic Net	0.36	0.16
Linear Regression	0.35	0.17

- **Results and Insights**

- year of release didn't affect the score at all it was expected as year of release has no correlation with any of the features as we saw in the data exploration part.
- combination of the given features the model more info to predict the sales with R2-Score 0.66 with removing outliers
- Keeping outliers of the features increase R2-Score to 0.87 but increased RMSE to 0.71 instead of 0.23
- Adding new features like Critic Count and User Value increased the score by 3%.
- Finally the model could predict sales with a good acceptable score.

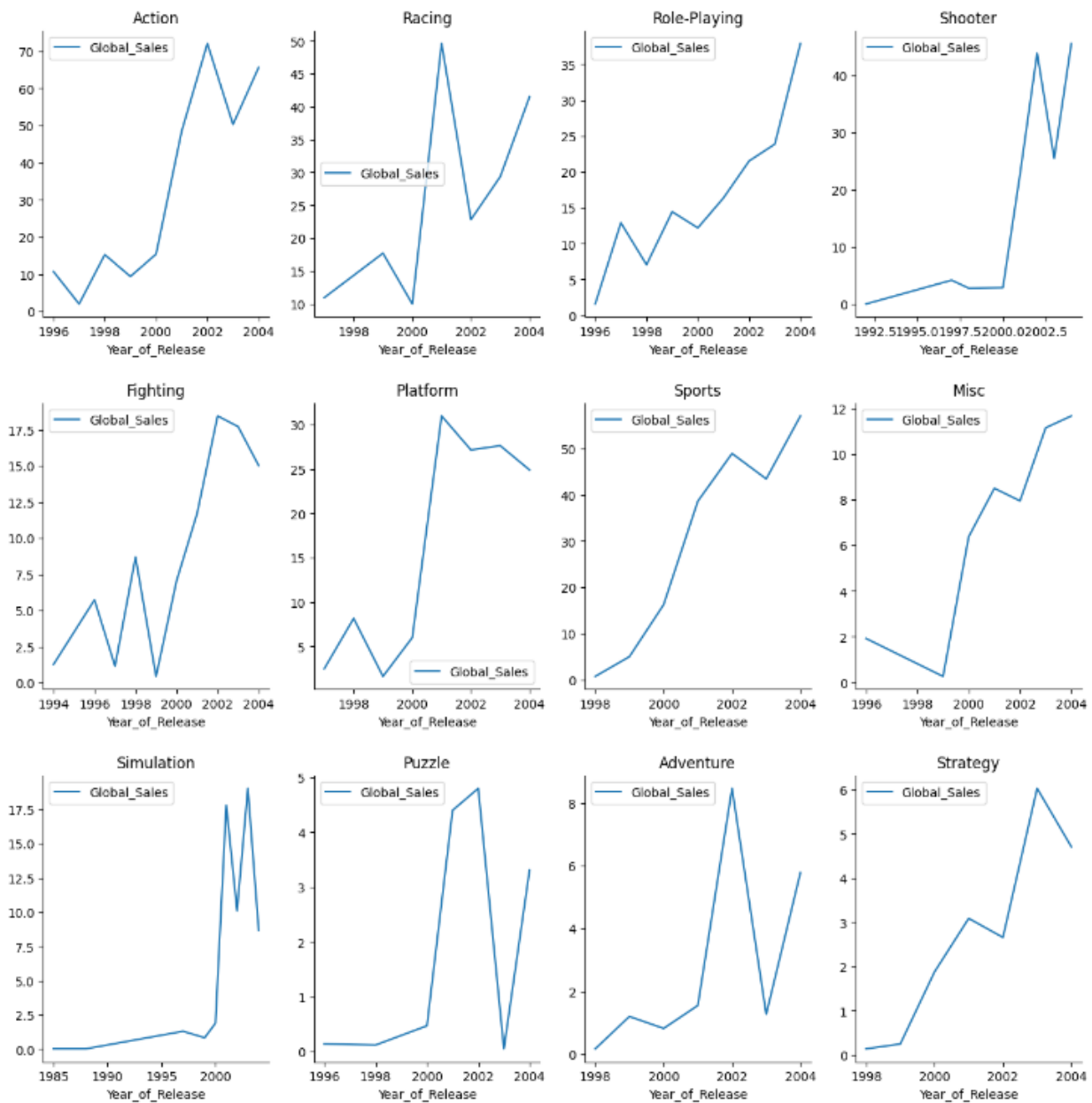
## Q2: For a specific genre , will its sales increase/decrease over the upcoming years?

- Expectations

genre that has high sales over previous years, its sales will increase over upcoming years unless, There's a great drop in some years that may affect the upcoming sales.

- EDA

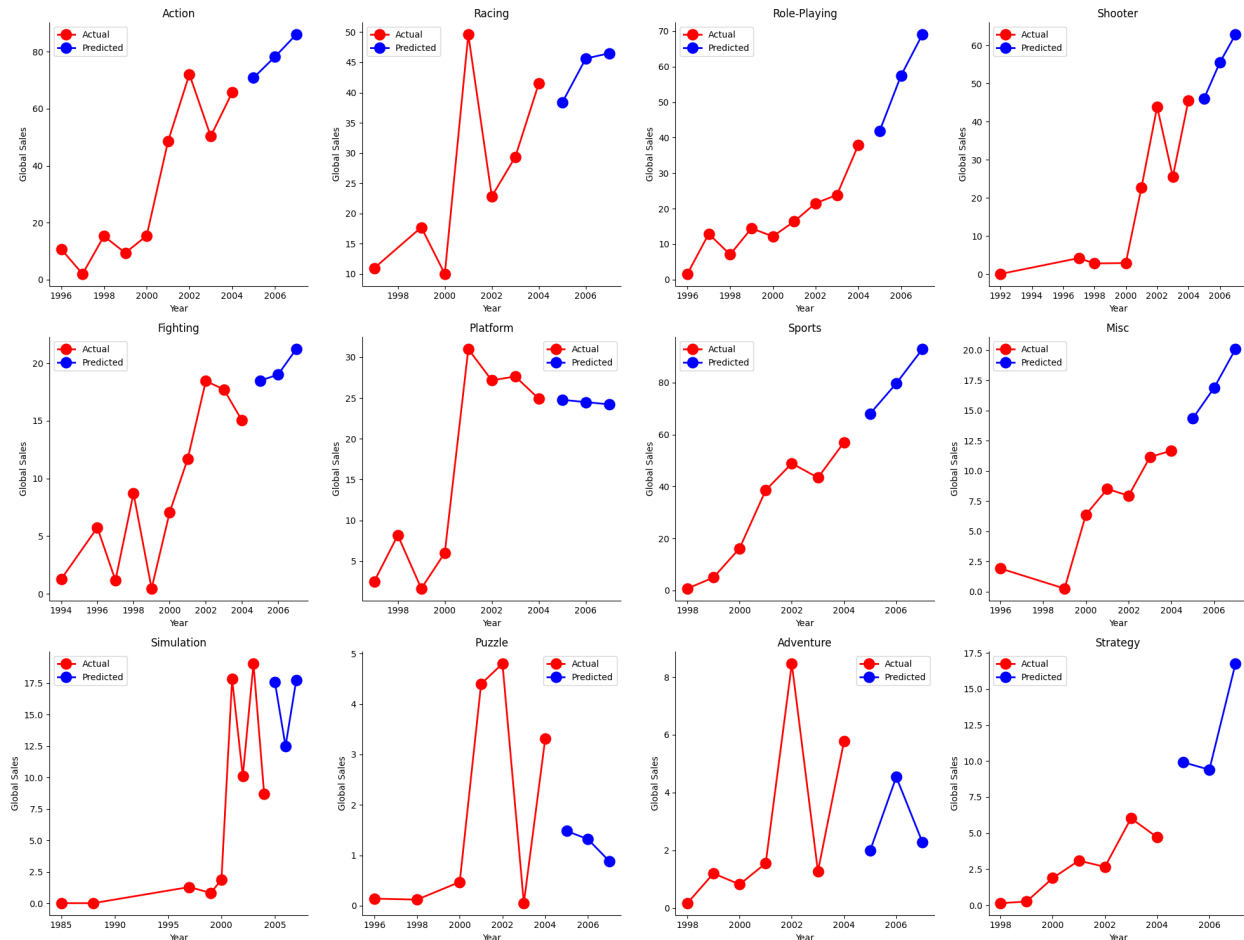
- Global Sales Vs Years for each genre



## ✓ Results

- Action, Racing, sports, role playing, shooting, puzzle and adventure sales are increasing.
- Simulation, strategy and fighting sales are decreasing.

### ● Time Series Prediction



### ● Results and Insights

- For the next 3 years the model predictions are:
- Action, Racing, sports, role playing, shoot sales will keep increasing.
- puzzle genre sales will decrease, that's because it has a great drop in its sales in 2002
- adventure genre sales will increase in the first year then decreases as it has a great drop also in 2002.

- Simulation sales will decrease in the first year, then decrease. because of the significant increase in 2000.
- Fighting genre, the same as simulation, will increase but its peak was from 2000 to 2002.
- So, Data matches the expectations

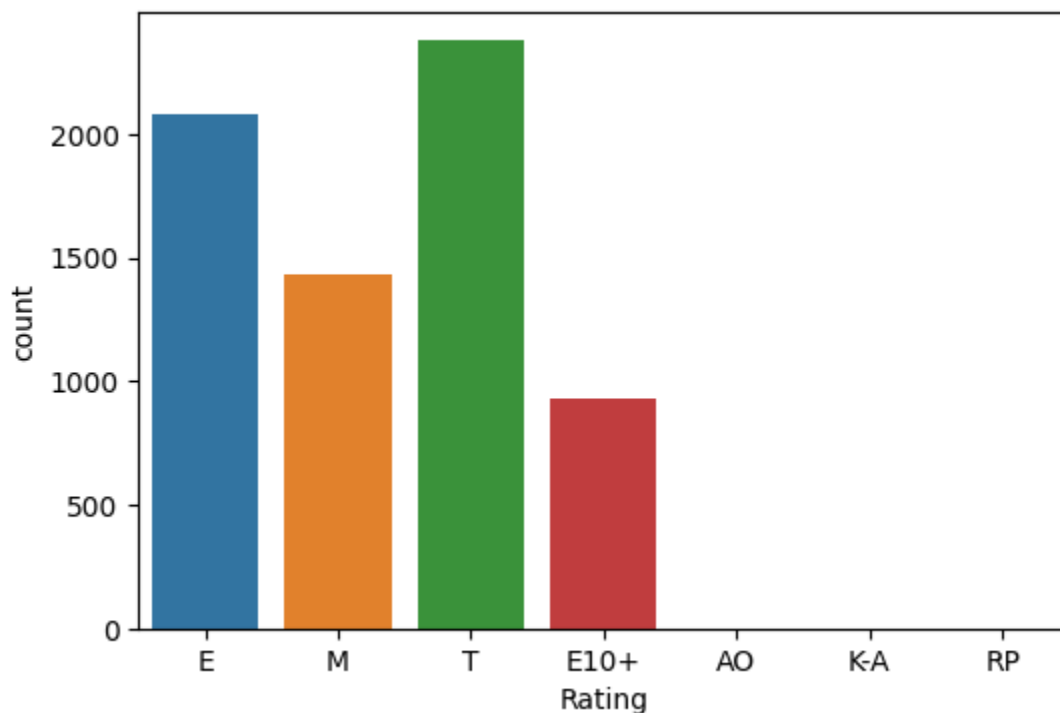
Q3: Can we predict Rating of a game based on features given: publisher, platform, critic score,... ? (EXTRA QUESTION)

- Expectations

- The model should determine the game characteristics through critic score, developer, publisher and platform that best predict that Rating of a game.

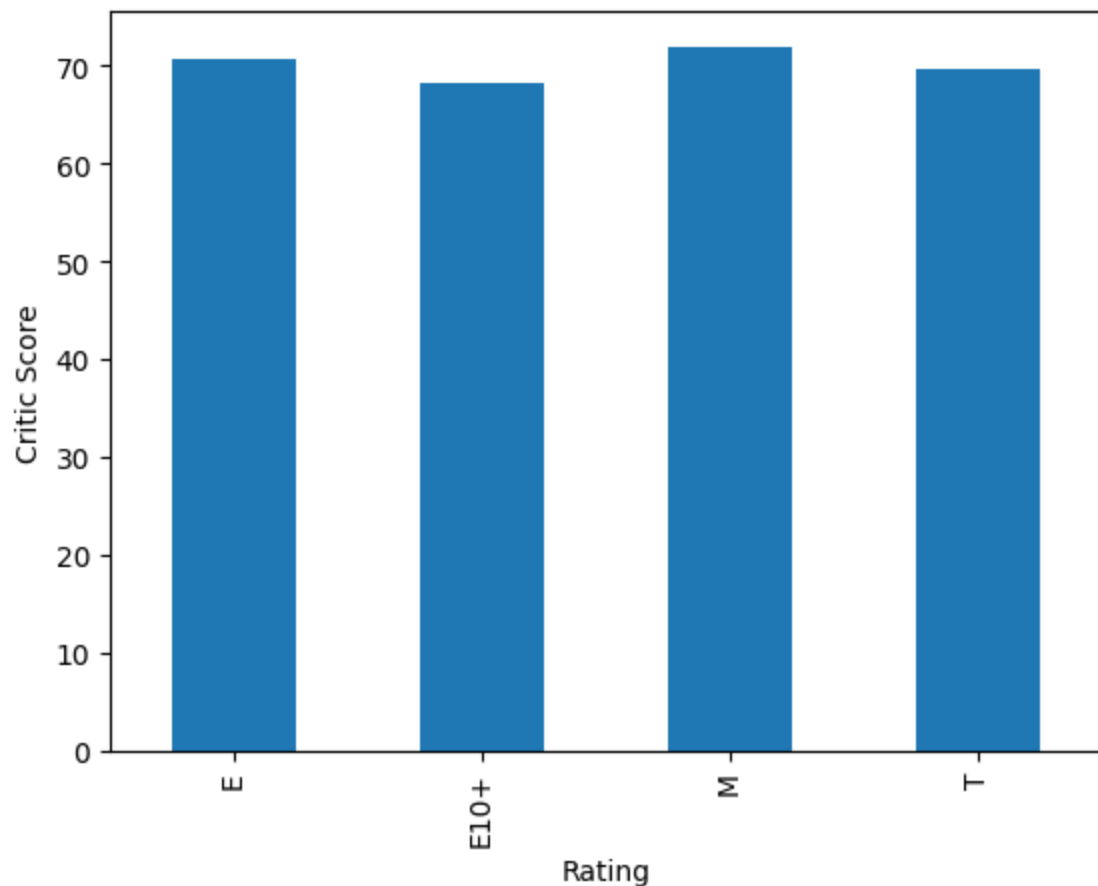
- EDA

- Rating values histogram shows that we have only 4 values in the dataset for rating
  - Rating in our data means this game suitable for which age

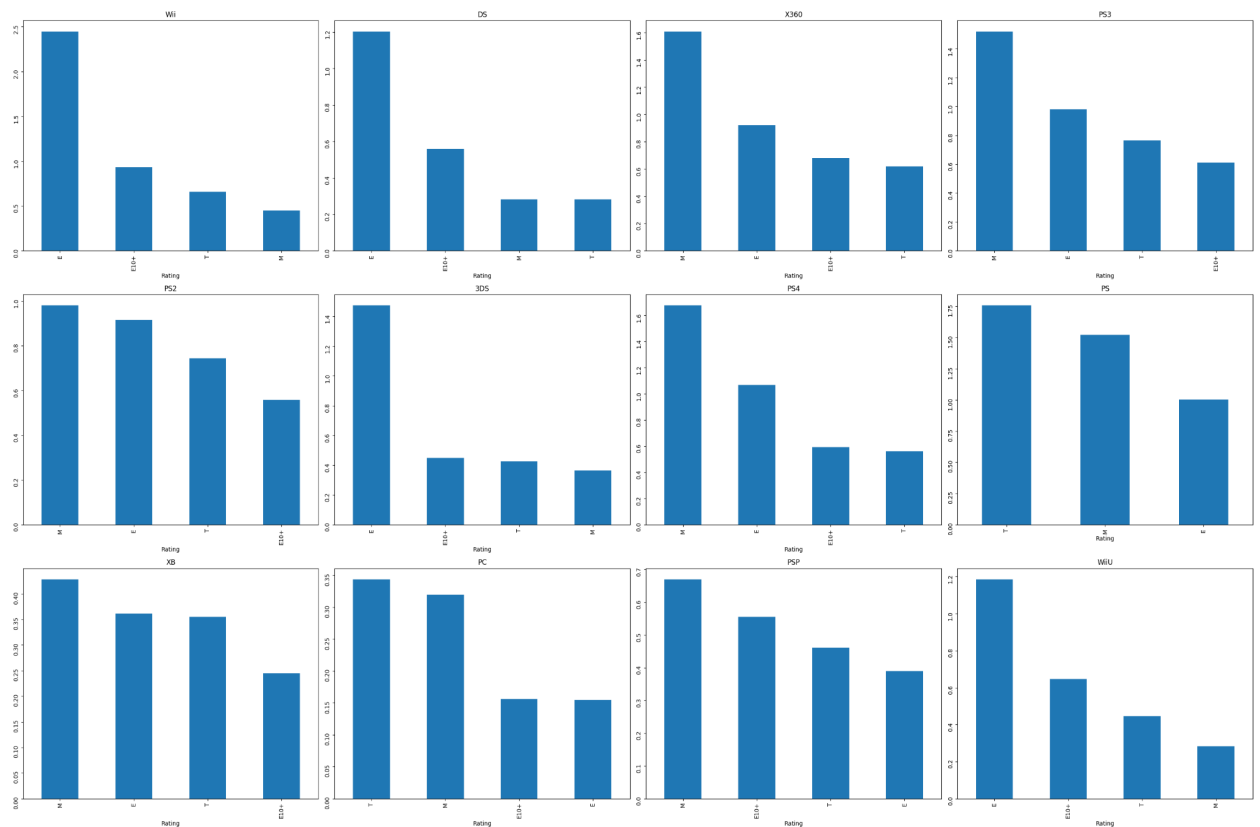


- Critic Score in each Rating class





- it seems values are very close so let's look at critic score for each rating w.r.t to another feature: for each Platform : Mean critic score for each Rating class, now we have variety so Platform will help discriminate between different classes
- **Global sales for each rating class:** it seems that sales is a good feature to discriminate between classes. We have significant differences between means of sales.



- Hypothesis Test

- H0: Mean Critic score the same for all rating class

- HA: each class has different mean of critic score

- ☐ Using ANOVA Test: P-value =

- 2.283678202155225e-10 < 0.05 so reject null hypothesis

- ☐ Critic score will help discriminate between rating classes

- Models

Model	Accuracy
SVM	80%
Random Forest	77.5%
LR	77%
Adaboost	76.5%

- Results and Insights:

Data match expectations: Model succeed in predicting Rating of a game given sales, critic score, publisher, developer ,platform.

## 5.4. Mechanistic Questions

Q1: How does the choice of platform affect sales for a specific genre?

- Expectations

Some platforms are more friendly to specific genres apart from others, the reason is that controllers / hardware of each platform is different and designed with specific type(s) of games in mind.

- EDA

Relevant columns: Platform, Genre and Global\_Sales

Nulls: Only two nulls in Genre, so we just filter those out

Outliers: Outliers in Global\_Sales get removed

Next step is exclude Platforms with very little number of games(<200), this ends up removing around 15 platforms

We then Group By Platform and Year\_of\_Release, then get the mean of sales.

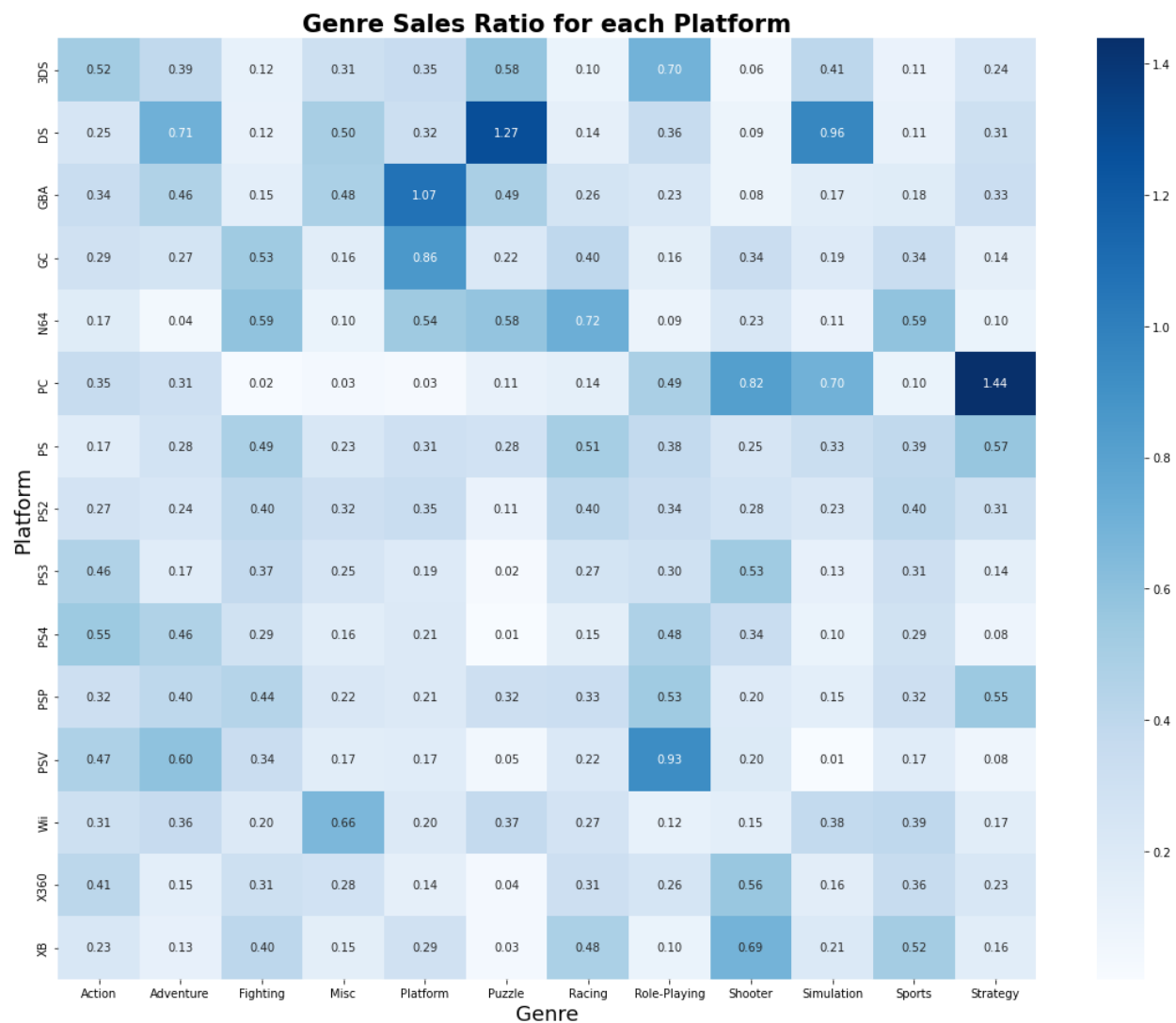
	Platform	Genre	Global_Sales
14	DS	Fighting	7.11
7	3DS	Role-Playing	16.63
174	XB	Racing	21.05
116	PS4	Shooter	6.49
95	PS2	Strategy	15.15
44	GC	Shooter	9.48
3	3DS	Misc	9.06

Here is a treemap of some of the results so far



Now, we for each genre in each platform we convert the Global\_Sales into a ratio, we divide it by total sales in that platform AND total sales of that genre for normalization purposes

Then we plot our results:



Top score to explain mechanic behind:

[('PC', 'Strategy', 1.4381), ('GBA', 'Platform', 1.069),  
 ('DS', 'Adventure', 0.7076), ('PSV', 'Role-Playing', 0.9252),  
 ('PC', 'Shooter', 0.8194),

- Results and Insights

We can clearly see that our explorative part of the question is done and the results matched the expectation, there are dominant genres in most platforms.

For the Mechanistic part, here is a few sample explanations:

### 1. PC & Strategy Games

- PC is a very very flexible platform and has a wide variety of setups, users basically create their dream platform using it.
- Strategy games often involve a lot of observations and focus on signals and metrics with too much detail.
- PC players can have a setup of more than one monitor for the details, observation, signals and metrics to watch out for.
- PC players also have mouse, keyboard and even joystick to offer both precise and fast controls

### 2. GBA & Platform

- GBA is a handheld console with D-pad and a very limited number of buttons, this makes it perfect for casual players who just want to plug-and-play with no hassle.
- Platform games are usually played in short sessions, have basic movements and not very much is going on the screen.
- Platform games could be fast-paced which requires fast controls.
- GBA also has a D-pad which is both fast to reach and precise to control.

## 6. Conclusion

### ☐ Benefits & Business Actions should be taken

- Company should deal with top publisher-stated below- to gain more global sales for a new game they want to launch
  1. Nintendo
  2. Microsoft Game Studios
  3. Take-Two Interactive
  4. Sony Computer Entertainment
  5. Electronic Arts
  6. Activision
  7. Warner Bros. Interactive Entertainment
  8. Square Enix
  9. Disney Interactive Studios
  10. Capcom
- Which game genre to publish and in which region?
  - If you want to publish a Sports' game, then you'd definitely go to NA and we can guarantee an average sales of 0.2 million dollars.
  - For a Shooter's game, go to Europe and we'd guarantee an average sales of 0.06 million dollars.
  - It's preferable to not publish a game in the Japan region, but if you want to try your luck, go for the Role-Playing genre and if you're lucky enough you'd have an average of 0.03 million dollars.
- Which of the platforms is stronger or more stable in terms of sales of games published to it? Explo. Q3
- What platforms / genres are more loved by Critics / Users? Explo. Q4
- Company should produce games in the following genres as its sales will be increased over the upcoming years which we predicted using our time series model:
  - Action, Racing, sports, role playing, shoot



- We offer the company a model to predict the global sales of a new game given its publisher, developer, platform , class rating (min. allowable age), and critic score
- We offer the company a model to predict the Rating class of a game given developer, publisher, platform and critic score.
- Which genre to publish for a specific platform? Mech. Q1
- Which platform to target for a specific genre? Mech. Q2
- Which regions to focus on? Expl. Q6
  - North america has max sales for most publishers
- Which genres to choose? Desc. Q1
  - Action then Sports