

VIDEO GAME

Yushu Gong
Ziqiao Zhang

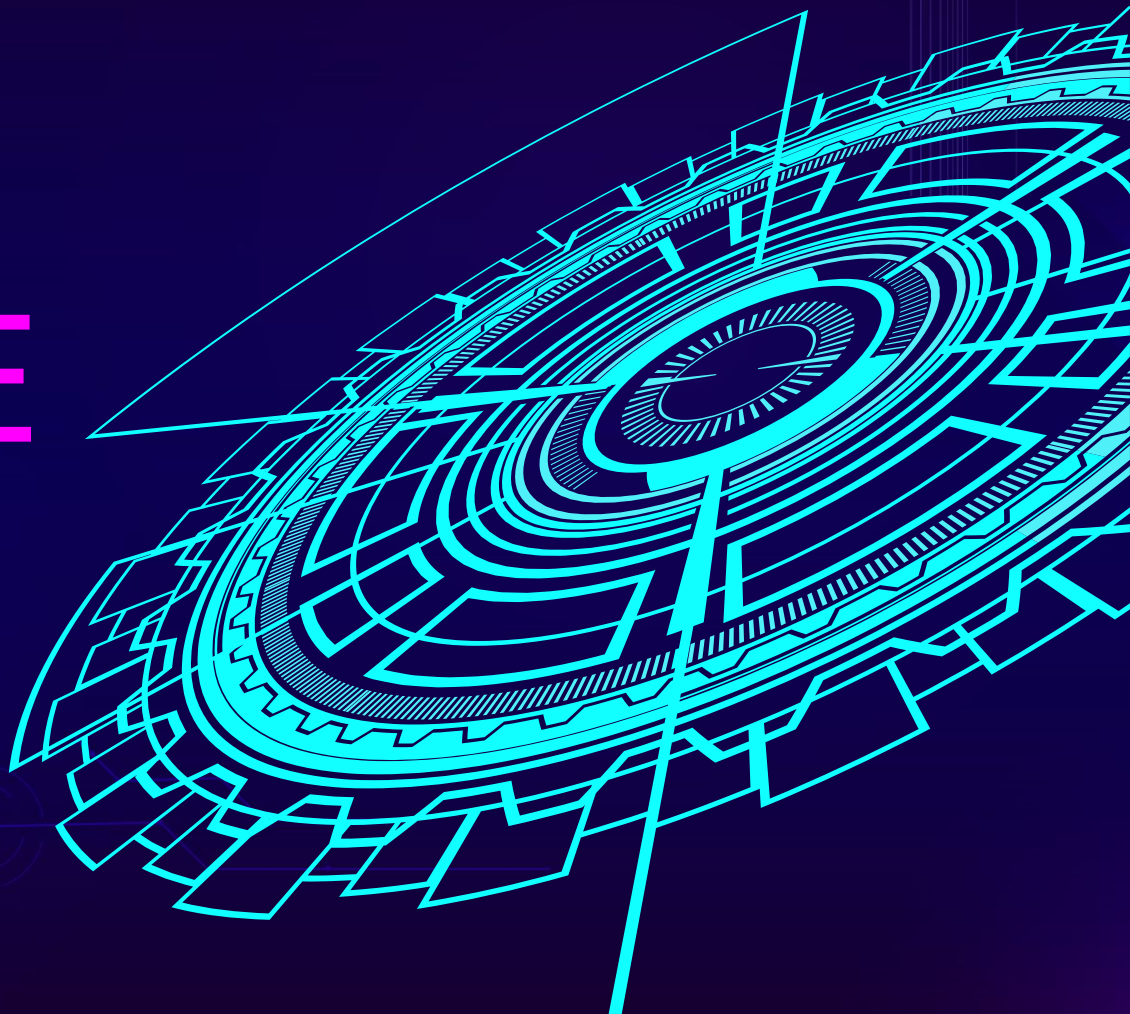




TABLE OF CONTENTS

01 Introduction

02 Platform & Genre

03 Sales & Genre

04 Limitations

INTRODUCTION

Video games, is a form of entertainment that creates interaction with the user, as well as visual effects. Through constant iterations of hardware and software, now became an important universal force that beyonds country borders and cultures.





Dataset & Clean

- In terms of data cleanliness, the dataset is well-organized. Just remove some null values and duplicate lines.
- This is a relatively new dataset. It includes the video games from 1970 - 2023. To improve completeness and visualisation, we selected the data from 1980 to 2022 in it.
- The problem is some of the columns' type in the dataset are 'list'. During analysis, the lists needed to be split into multiple lines of 'str' type and eliminate extra punctuation.

	Title	Release Date	Developers	Summary	Platforms	Genres	Rating	Plays	Playing	Backlogs	Wishlist	Lists	Reviews
0	Elden Ring	Feb 25, 2022	['FromSoftware', 'Bandai Namco Entertainment']	Elden Ring is a fantasy, action and open world...	['Windows PC', 'PlayStation 4', 'Xbox One', 'P...	['Adventure', 'RPG']	4.5	21K	4.1K	5.6K	5.5K	4.6K	3K
1	The Legend of Zelda: Breath of the Wild	Mar 03, 2017	['Nintendo', 'Nintendo EPD Production Group No...']	The Legend of Zelda: Breath of the Wild is the...	['Wii U', 'Nintendo Switch']	['Adventure', 'Puzzle']	4.4	35K	3.1K	5.6K	3K	5.1K	3K
2	Hades	Dec 07, 2018	['Supergiant Games']	A rogue-lite hack and slash dungeon crawler in...	['Windows PC', 'Mac', 'PlayStation 4', 'Xbox O...']	['Adventure', 'Brawler', 'Indie', 'RPG']	4.3	25K	3.5K	7.3K	4K	3.2K	2.1K
3	Hollow Knight	Feb 24, 2017	['Team Cherry']	A 2D metroidvania with an emphasis on close co...	['Windows PC', 'Mac', 'Linux', 'Nintendo Switch']	['Adventure', 'Indie', 'Platform']	4.4	25K	2.7K	9.6K	2.6K	3.4K	2.1K

Dataset & Clean



- The dataset is organized with a high level of cleanliness
- To ensure its integrity, addressing unknown values and eliminating
- The primary emphasis of data collection spans from 1980 to 2020, but has fewer games post-2016. To maintain the consistency of our analysis, we opted to exclude data recorded after 2016.

Name		Platform	Year	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales
0	Wii Sports	Wii	2006	Sports	Nintendo	41.49	29.02	3.79	3.31	35.82
1	Super Mario Bros.	NES	1985	Platform	Nintendo	29.08	3.58	3.28	2.96	33.00
2	Mario Kart Wii	Wii	2008	Racing	Nintendo	15.85	12.88	10.22	1.00	31.37
3	Wii Sports Resort	Wii	2009	Sports	Nintendo	15.75	11.01	3.28	2.96	33.00
4	Pokemon Red/Pokemon Blue	GB	1996	Role-Playing	Nintendo	11.27	8.89	10.22	1.00	31.37

Year		Game_Count
0	Adventure	2
1	2020	1
2	2017	3
3	2016	344
4	2015	614
5	2014	582
6	2013	546
7	2012	657
8	2011	1139
9	2010	1258



Research Questions

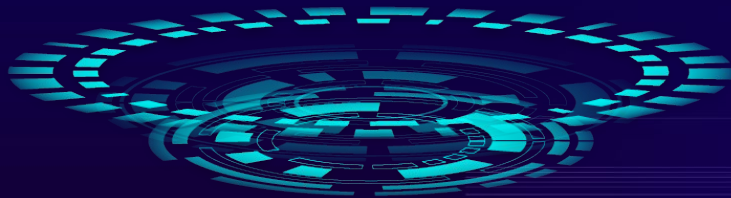
The trend of **genres** and **platforms** over years. What is the **relationship** between them?

What is the relationship between regional and global game **sales**, and how do **genres** impact on global **sales**?



Platform & Genre

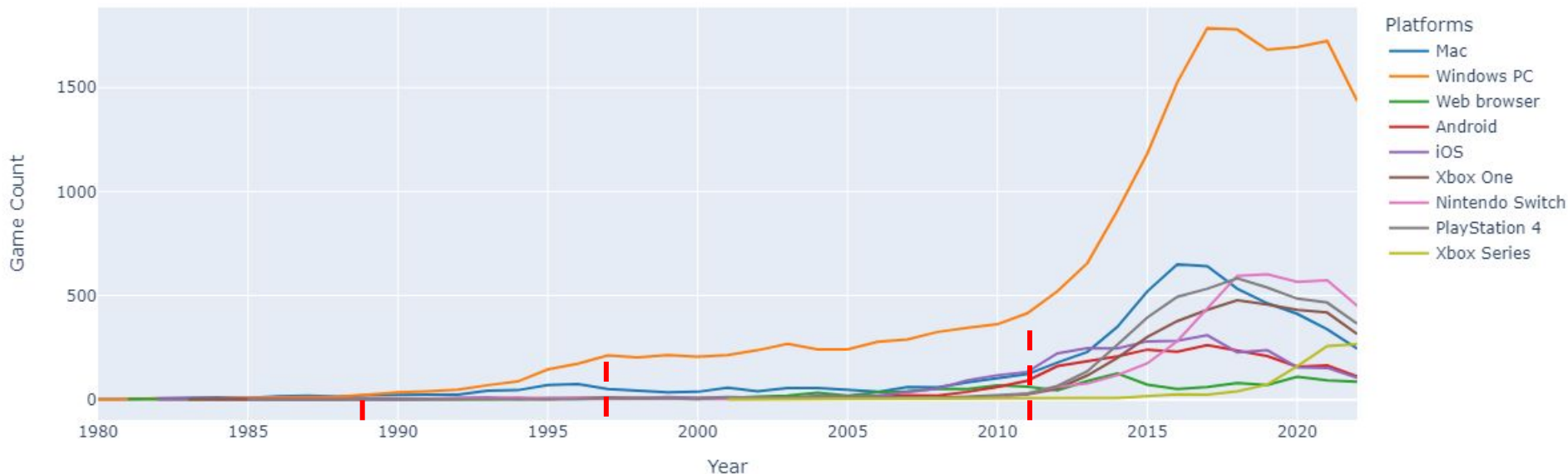
The trend of genres and platforms over years.
What is the relationship between them?





Platforms Trend

Main Platforms And Release Year



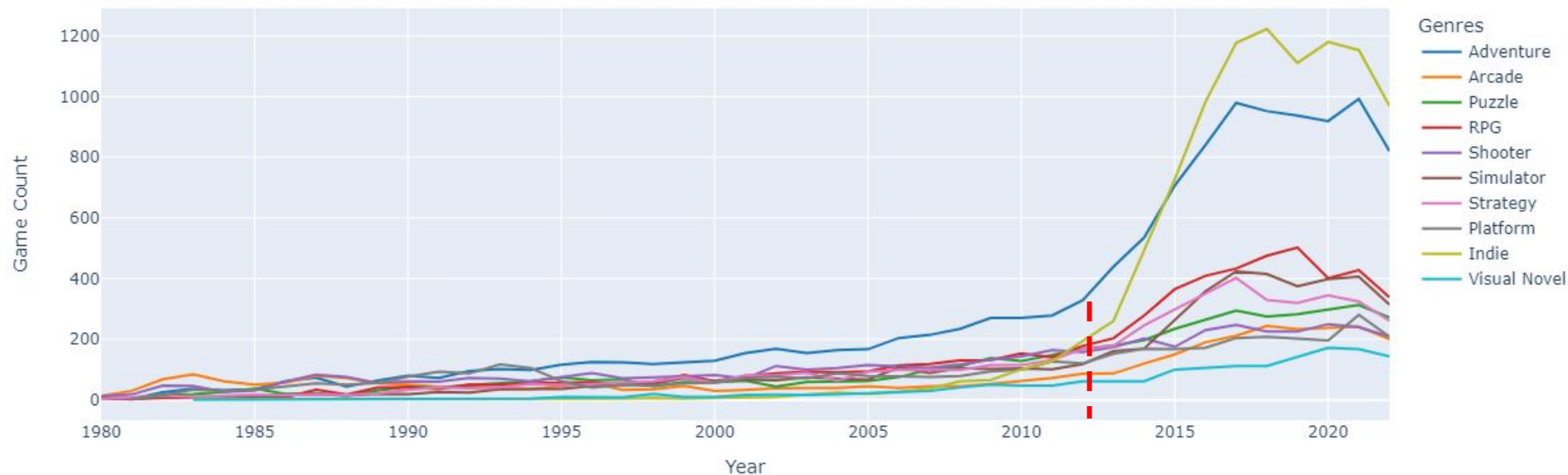
The evolution of the **Windows PC** is very obvious. 1988: started to emerge and grew rapidly.

1997: entered a steady development period. 2011: grew again with a sharp increase.



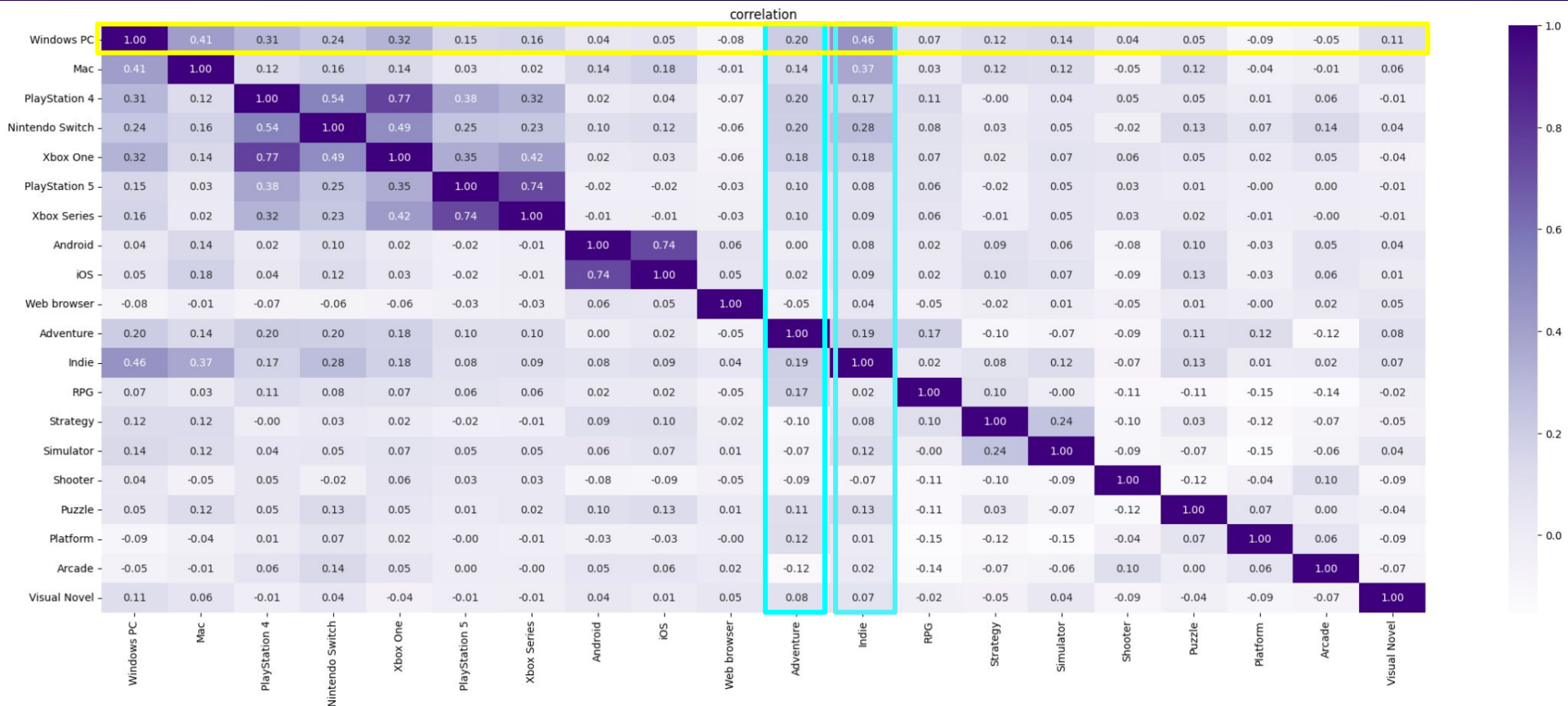
Genres Trend

Main Genres And Release Year



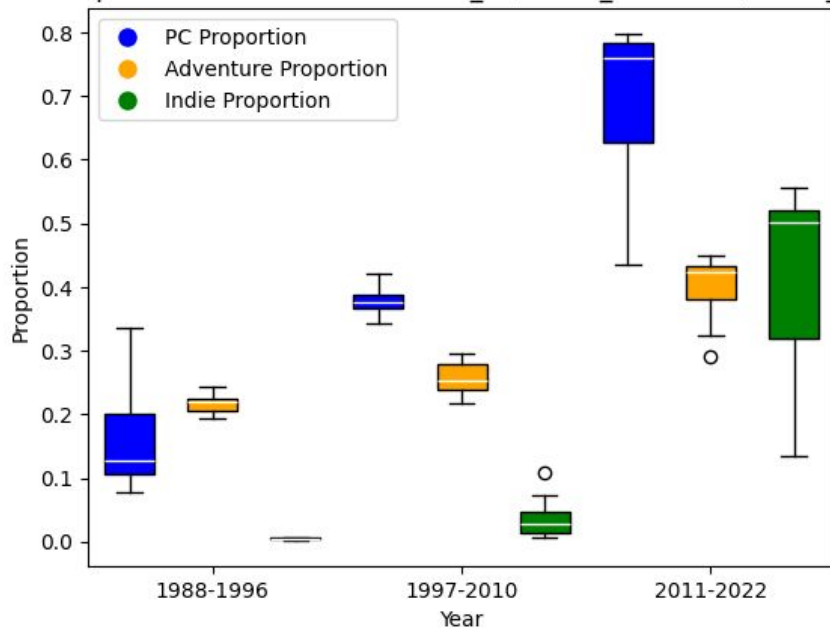
Indie and **adventure** are the two categories that currently hold the most market share. Adventure has a distinct development since 1995. Both of them has a sharp increase around 2012

Heatmap



Boxplot of Windows_PC/AVG/Indie games proportion

Boxplots for Proportions of Games in Platform_PC/Genre_Adventure/Genre_Indie by Period



Median

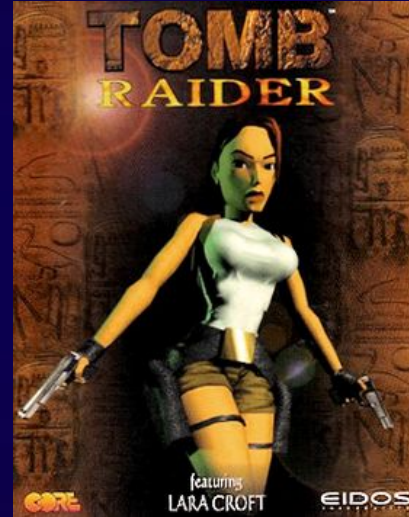
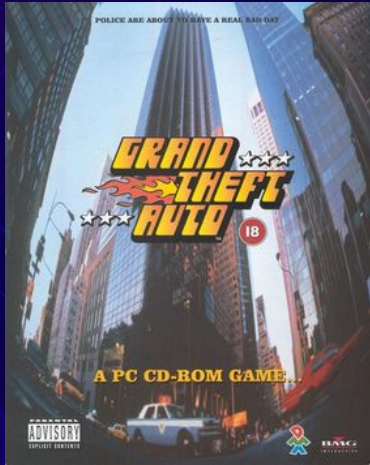
PC:
12% to 38% to 77%

AVG:
22% to 25% to 43%

Indie:
0% to 4% to 49%

Time Point

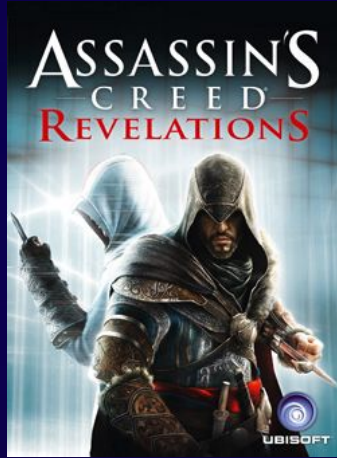
Around 1997



Universal access to 32-bit computers and 3D display cards
3D computer graphics
Unreal Engine

Time Point

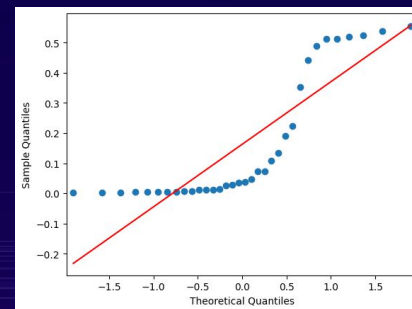
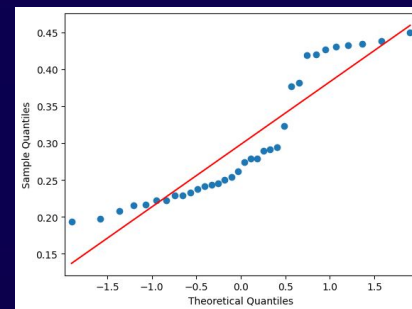
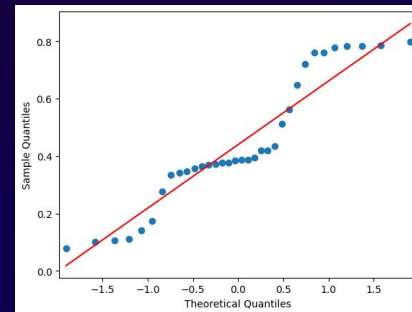
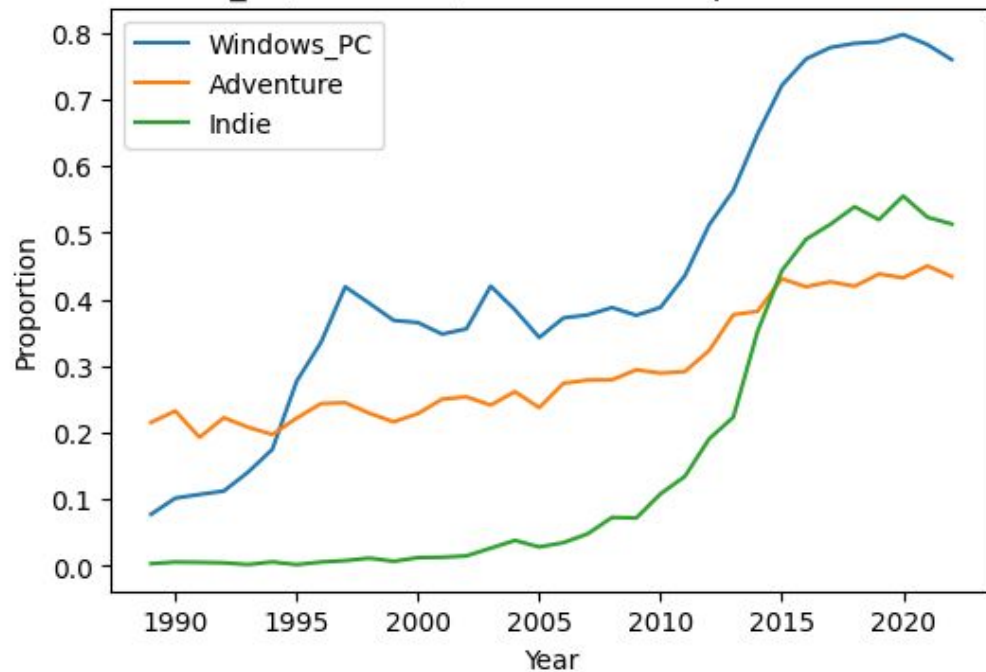
Around 2011



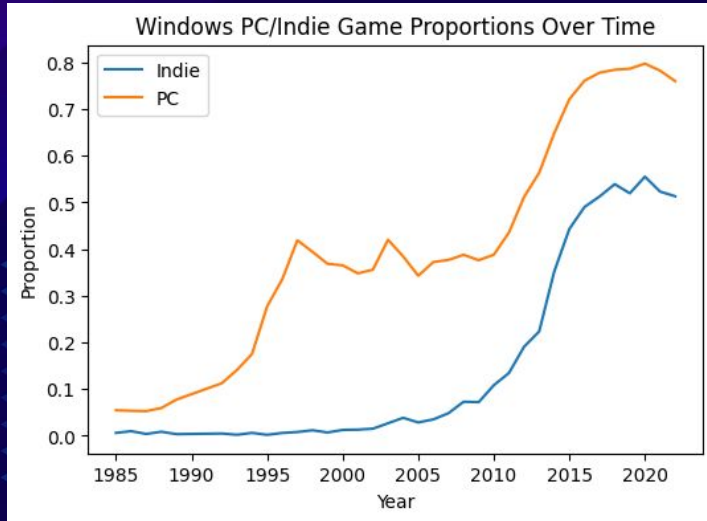
Mature game engine: Unity 3D Unreal4
Visual programming: Construct 2 GameMaker Pro
Attracting lots of indie game makers

Data Distribution

Windows_PC/Adventure/Indie Game Proportions Over Time



Spearman

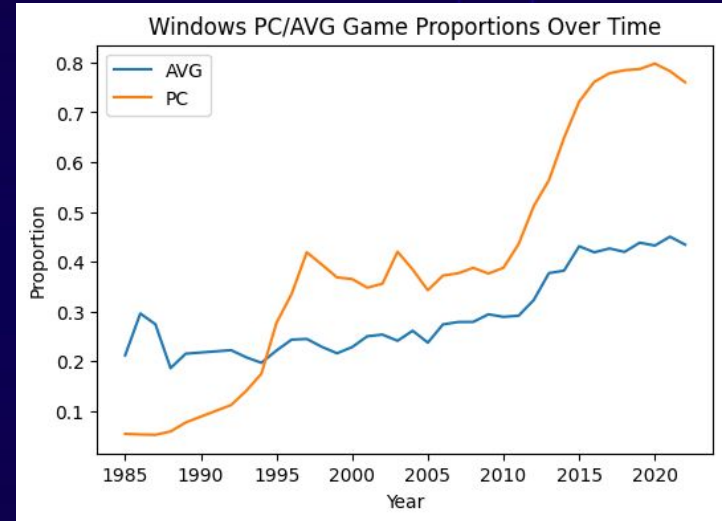


Windows PC & Indie

Spearman Rank Correlation: 0.92

P-value: 4.04e-15

significant monotonic relationship



Windows PC & AVG

Spearman Rank Correlation: 0.89

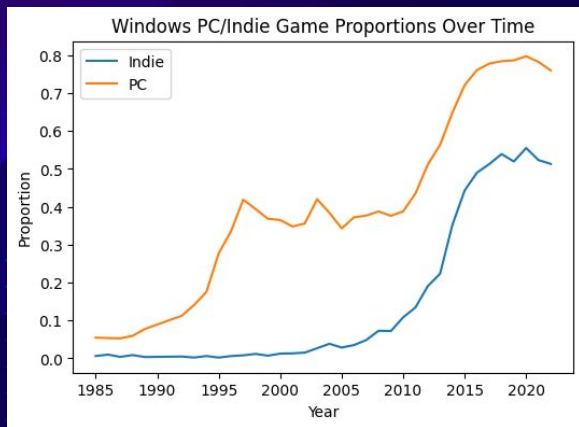
P-value: 1.18e-12

significant monotonic relationship

Heatmap



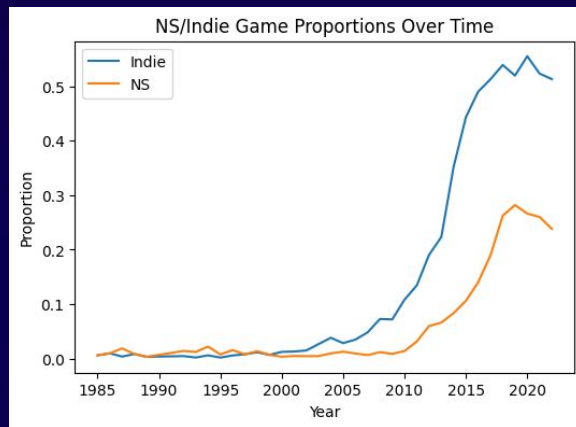
Indie



Windows PC & Indie

1988-2022
Spearman Rank Correlation: 0.92
P-value: 6.5e-06

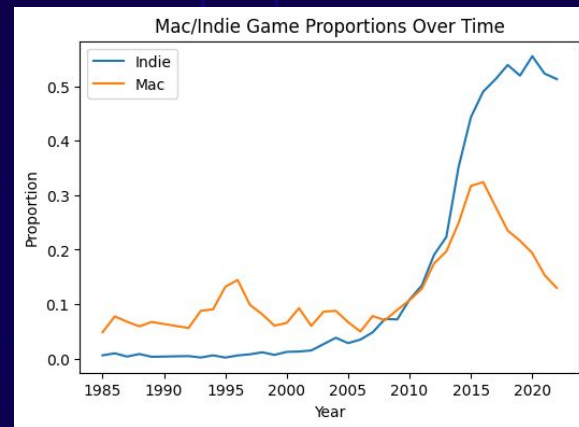
2011-2022
Spearman Rank Correlation: 0.95
P-value: 9.54e-07



Nintendo Switch & Indie

1988-2022
Spearman Rank Correlation: 0.67
P-value: 6.5e-06

2011-2022
Spearman Rank Correlation: 0.95
P-value: 9.54e-07

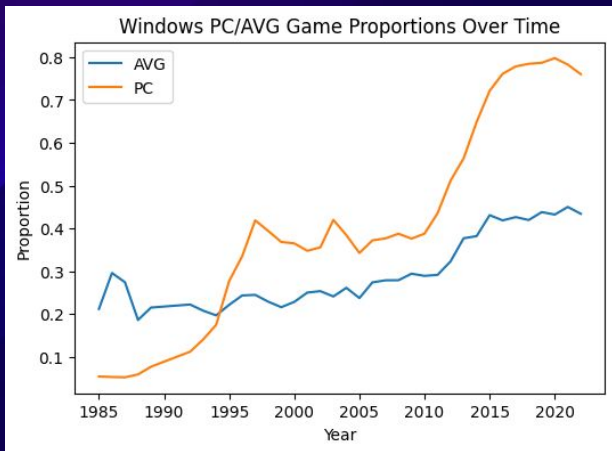


Mac & Indie

1988-2022
Spearman Rank Correlation: 0.64
P-value: 2.4e-05

2011-2022
Spearman Rank Correlation: 0.07
P-value: 0.82 >0.05

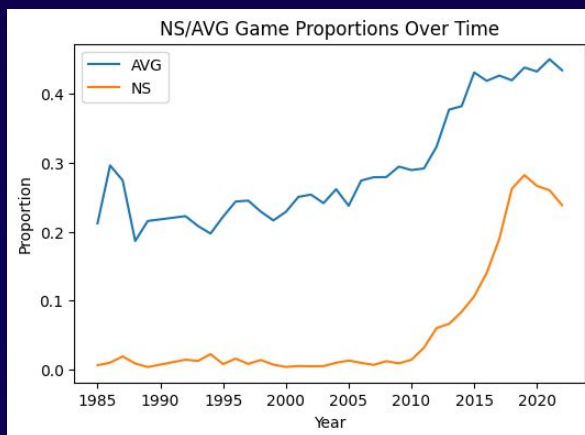
Adventure



Windows PS & AVG

1988-2022
Spearman Rank Correlation: 0.89
P-value: $5.3e-07$

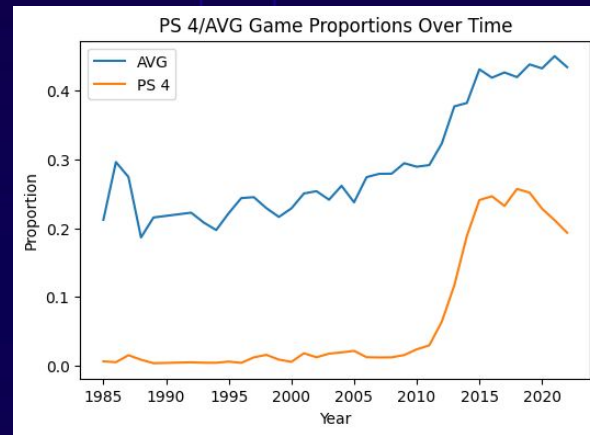
2011-2022
Spearman Rank Correlation: 0.77
P-value: 0.009



Nintendo Switch & AVG

1988-2022
Spearman Rank Correlation: 0.72
P-value: $5.3e-07$

2011-2022
Spearman Rank Correlation: 0.85
P-value: 0.0005



PS 4 & AVG

1988-2022
Spearman Rank Correlation: 0.83
P-value: $2.67e-10$

2011-2022
Spearman Rank Correlation: 0.55
P-value: $0.06 > 0.05$

conclusions

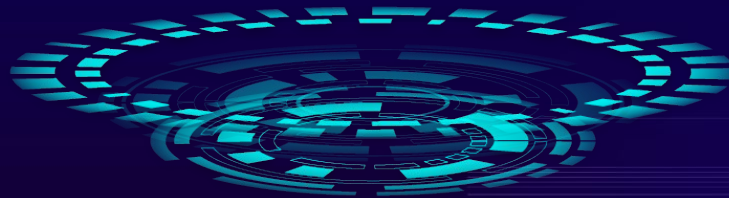
- Taking adventure games and indie games as examples, we can see that there is truly a **correlation** between different game platforms and game genres.
- But the impact is **changing**.

In the case of adventure and indie games, although the line graph shows that PC is rising very rapidly, and there is a strong similarity to the rising trend of AVG and Indie. But the extent to which it has influenced each game genre has changed at different times.



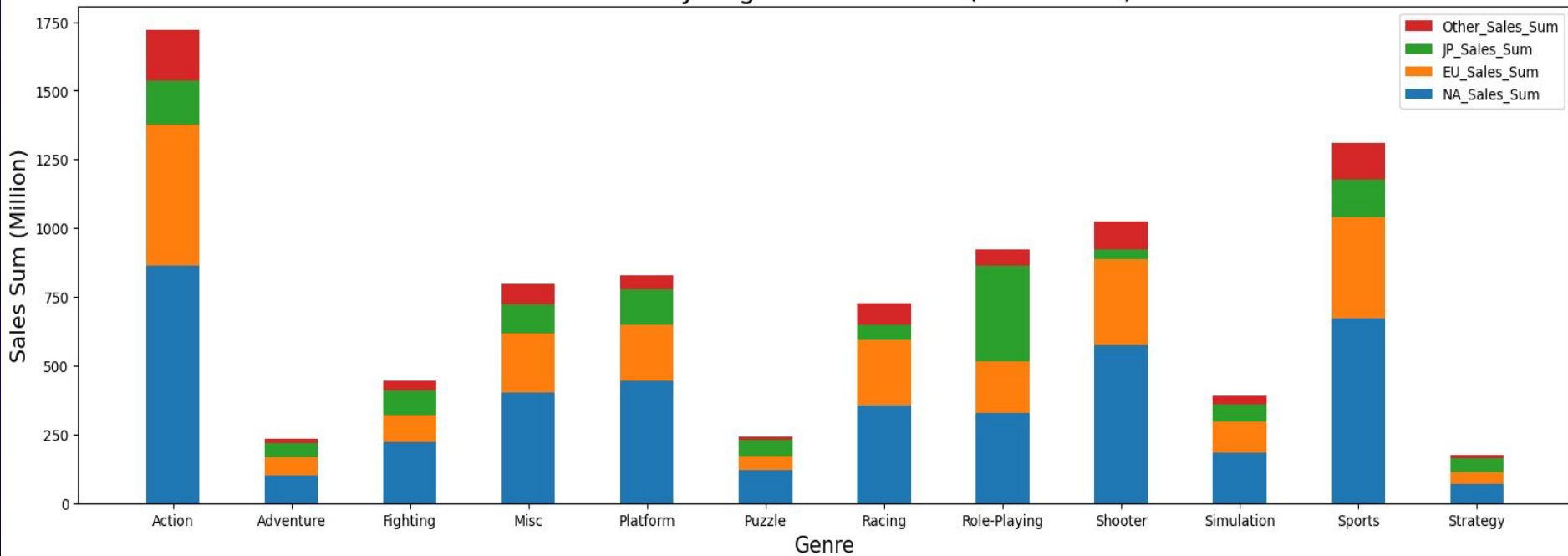
Sales & Genres

The Relationship Between Regions Sales and Global Sales
How Genres Affect Sales
Significance test



Region Sales & Global Sales

Sales Sum by Regions and Genre (1980-2016)



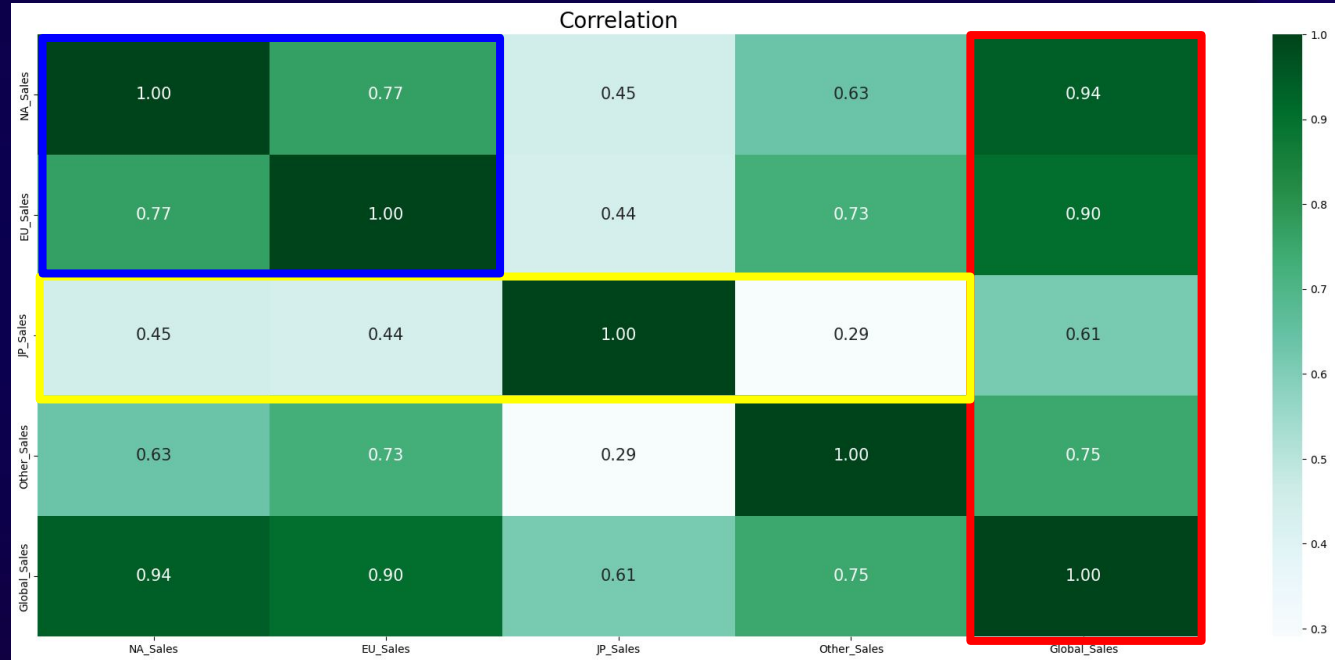
Region Sales & Global Sales

Why NA and EU have this close correlation?

The general reasons are contain: substantial populations, linguistic alignment and cultural impact.

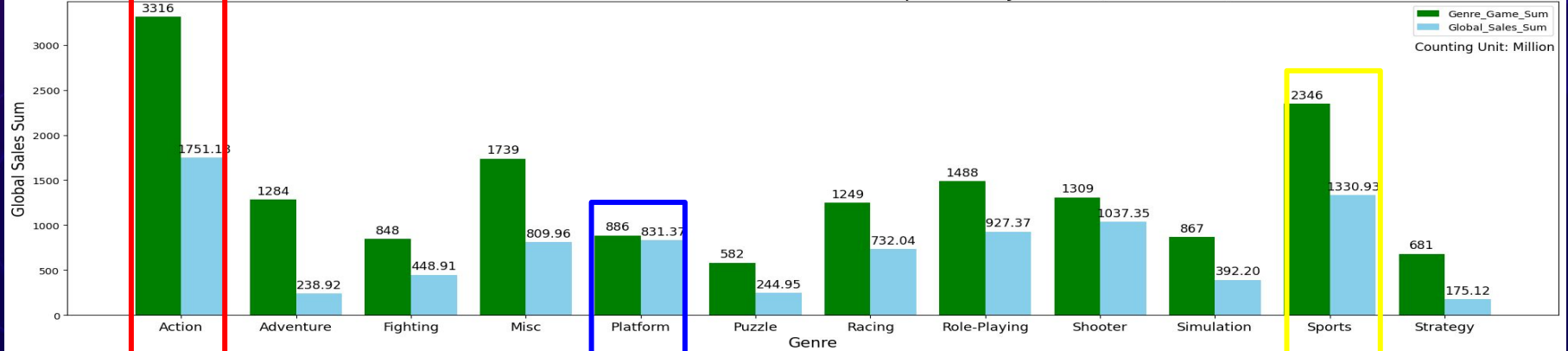
Why JP have less correlation with other regions?

The main reasons are:
Cultural Specificity and
Language differences.

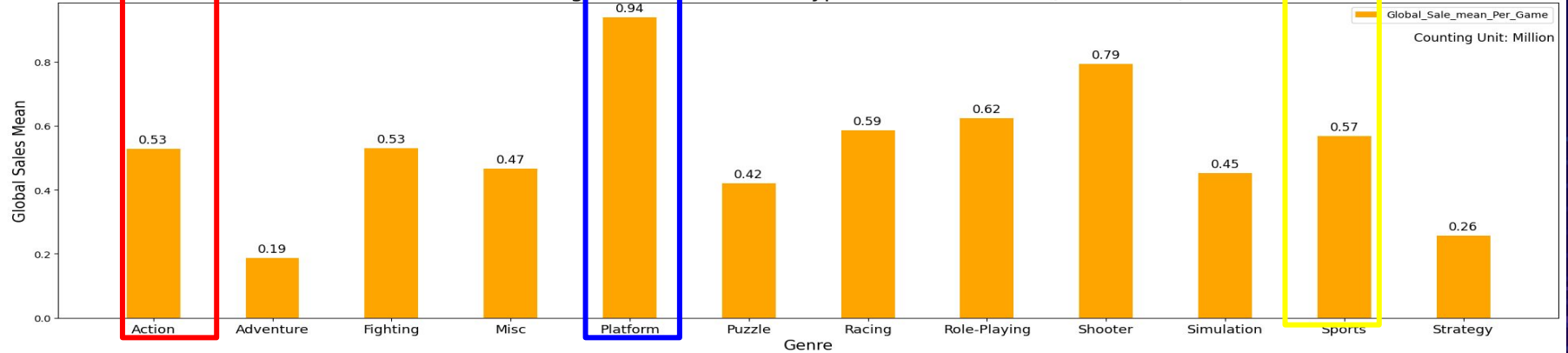


Genres & Sales

Global Video Game Distribution and Sales Comparison by Genre (1980-2016)

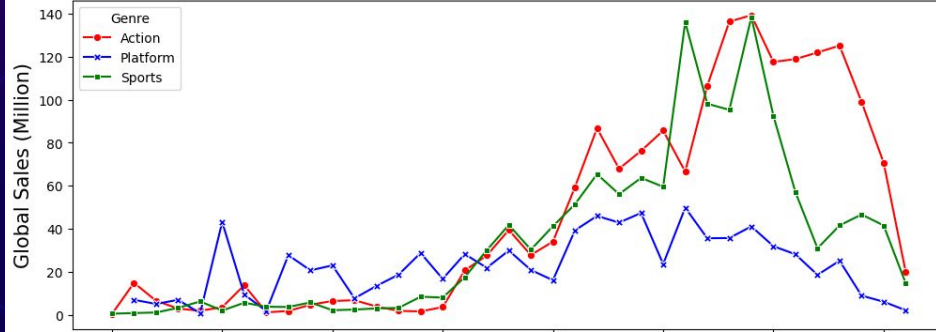


Global Average Sales of Different Types of Video Games (1980-2016)

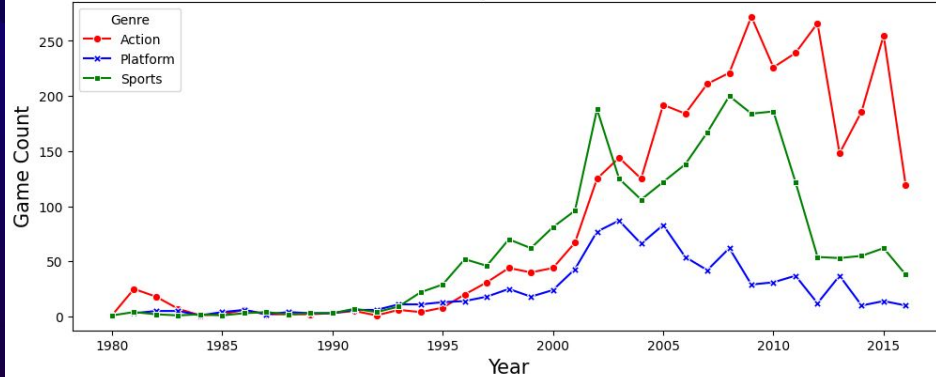


Genres & Sales

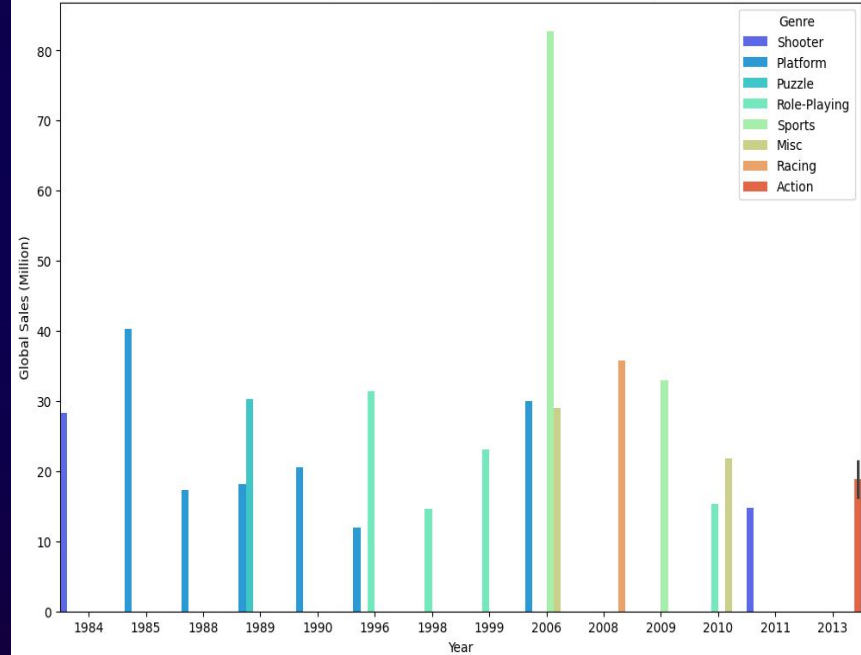
Global Sales for Action, Platform, Sports, and Shooter Genres (1980-2016)



Game Count for Action, Platform, Sports, and Shooter Genres (1980-2016)

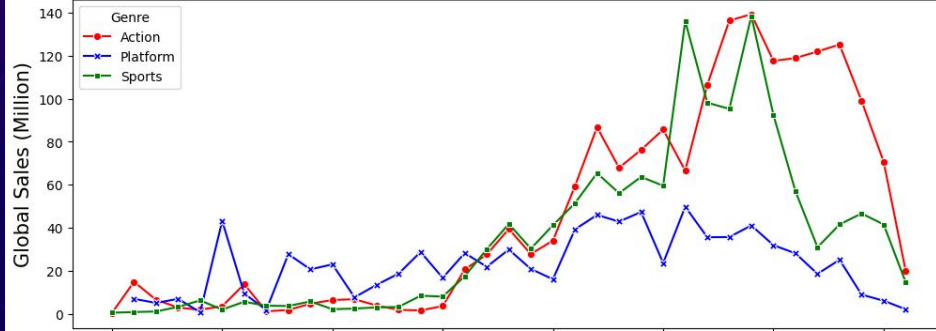


A Selection of the Top 5 Best-Selling Game Genres of Per Decade (1980-2016)

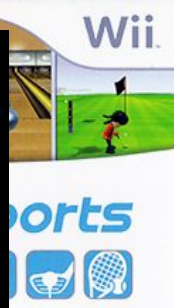
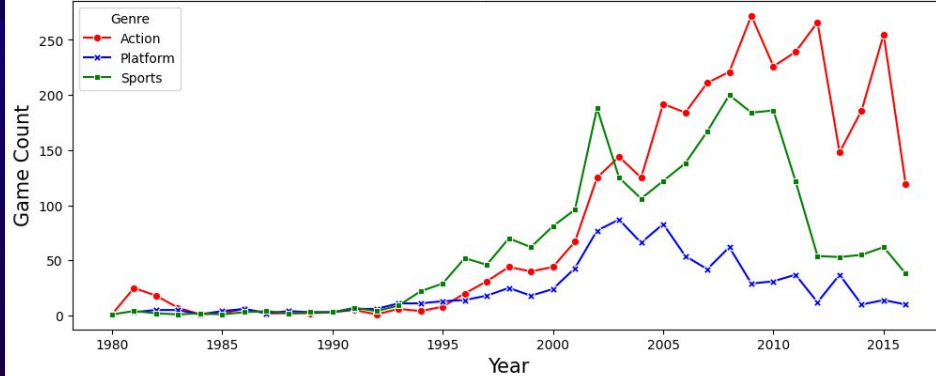


Genres & Sales

Global Sales for Action, Platform, Sports, and Shooter Genres (1980-2016)



Game Count for Action, Platform, Sports, and Shooter Genres (1980-2016)



Genres & Sales



Why have this trend?

Platform Game:

- Iconic franchises, fostering a loyal fan base.
- Easy gameplay mechanics, easily accessible to a broad audience.

Sports Game:

- Saturation of similar sports-themed games, leading to a lack of innovation

Action Game:

- Technological advancements, enabled the creation of more immersive action games.
- Saturated market with intense competition.
- Complex development processes, high production values, and extensive marketing campaigns.



- Spearman test to demonstrate whether there is a relationship between genres and sales.

Significant Test

Spearman Test	
P-Value	Effect Size
1.477e-09	0.047

Conclusions



The gaming industry's success is heavily influenced by the **North American market**.

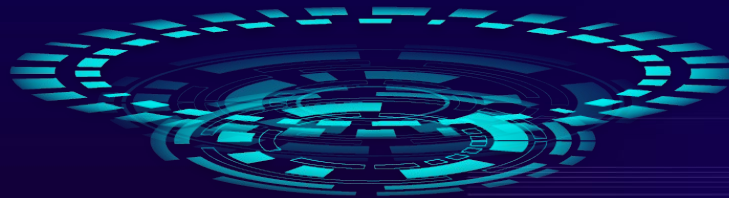
Action games are recently **popular genre**, given the **intense market competition**, the average sales in a moderate range.

Game genres can affect the global sales, but the effect size is quite small.



Limitations

This project has provided us with some insights.
But it's crucial to recognize and address specific limitations.



Limitations

- The dataset lacks comprehensive description, leaving uncertainty regarding how game genres were determined.
- Given the rapidly evolving nature of the gaming industry, the dataset may not capture the latest trends.
- Incomplete segmentation of the dataset's game genres.
- The platform data in the dataset only reflects the relationship between games released on that platform and the game genres. It cannot represent the relationship between the hardware devices and game genres.
- Acknowledging a limited proficiency in data science.



Thanks For Listening

