

https://github.com/RoisinMcG2022/UCDPA_RoisinMcGuire

In this report, I hope to offer some insight into the booming Gaming Industry. I will explore datasets on three main areas:

- 1: Platform options and popularity.
- 2: Genres, how tastes have evolved over time.
- 3: Streaming, in particular the impact of Twitch on the Gaming Industry.

Kaggle Datasets

I have drawn from the following four datasets to compile this report:

<https://www.kaggle.com/datasets/gregorut/videogamesales> vgsales.csv

<https://www.kaggle.com/datasets/rankirsh/evolution-of-top-games-on-twitch> Twitch_game_data.csv

<https://www.kaggle.com/datasets/atharvaingle/video-games-dataset> Video_Games_Dataset.csv

<https://www.kaggle.com/datasets/stetsondone/video-game-sales-by-genre> year_genre_region.csv

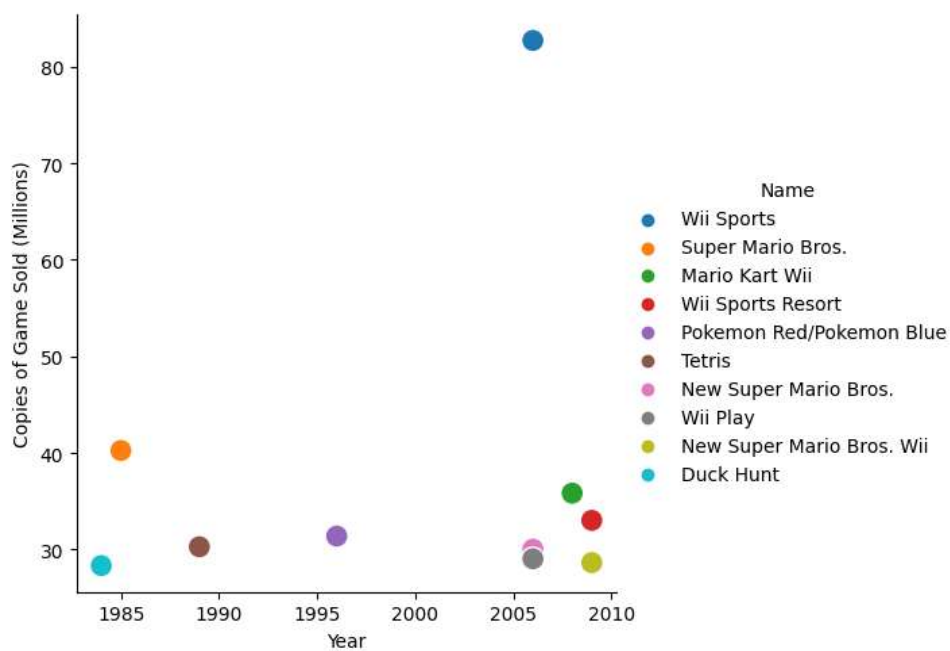
These datasets have been chosen because they cover the areas of interest as listed in the above Abstract. I would have preferred to use slightly more current data as the gaming industry has evolved very rapidly over the last five years but the listed options are free to access and validate. The chosen assortment of datasets can offer significant insight into how games have evolved from a childish pastime into an import social tool.

Throughout this report I will be using the word 'Platform' (Uppercase P) to refer to the game type, such as Super Mario Bros or Little Big Planet. I will use platform (lowercase P) to refer more broadly to gaming platforms, such as consoles, PCs and tablets.

Introduction

To begin this report, we shall look at the top ten of the bestselling Video Games between 1980 and 2016. (Dataset: vgsales.csv). There is one clear unifying factor between the top ten games and that is the publisher, Nintendo.

Rank		Name	Genre	Year	Publisher
0	1	Wii Sports	Sports	2006.0	Nintendo
1	2	Super Mario Bros.	Platform	1985.0	Nintendo
2	3	Mario Kart Wii	Racing	2008.0	Nintendo
3	4	Wii Sports Resort	Sports	2009.0	Nintendo
4	5	Pokemon Red/Pokemon Blue	Role-Playing	1996.0	Nintendo
5	6	Tetris	Puzzle	1989.0	Nintendo
6	7	New Super Mario Bros.	Platform	2006.0	Nintendo
7	8	Wii Play	Misc	2006.0	Nintendo
8	9	New Super Mario Bros. Wii	Platform	2009.0	Nintendo
9	10	Duck Hunt	Shooter	1984.0	Nintendo

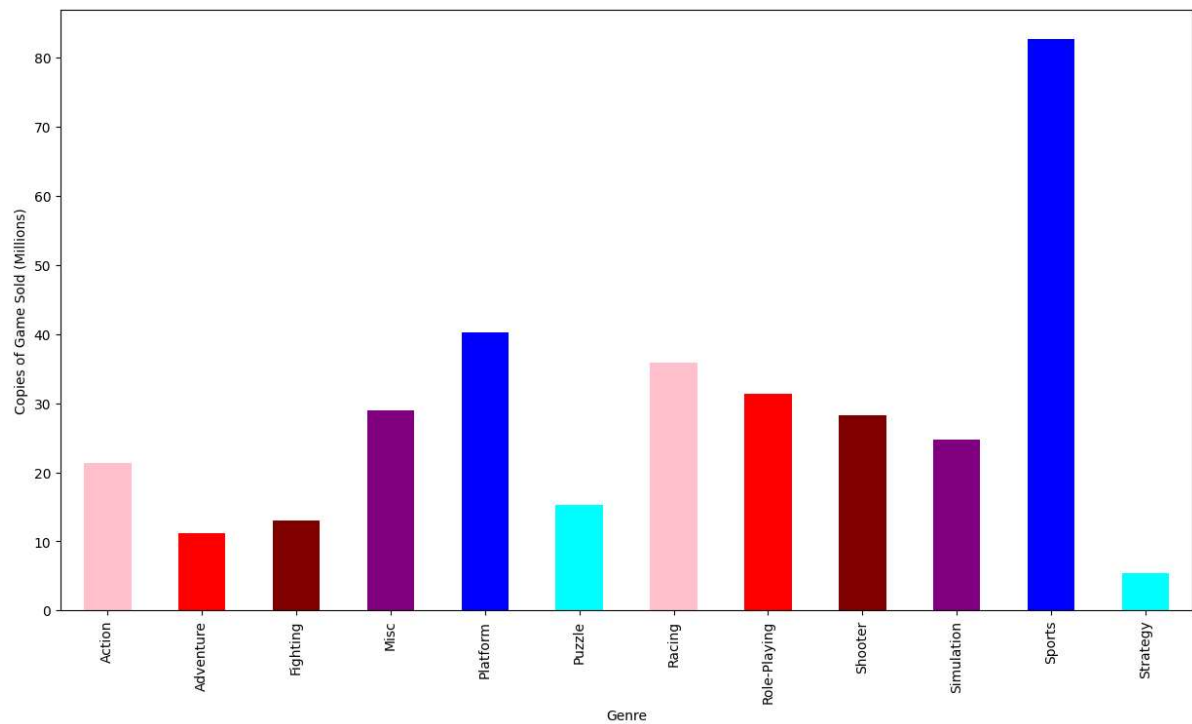


Nintendo's domination of the top ten bestselling games is clearly visible.

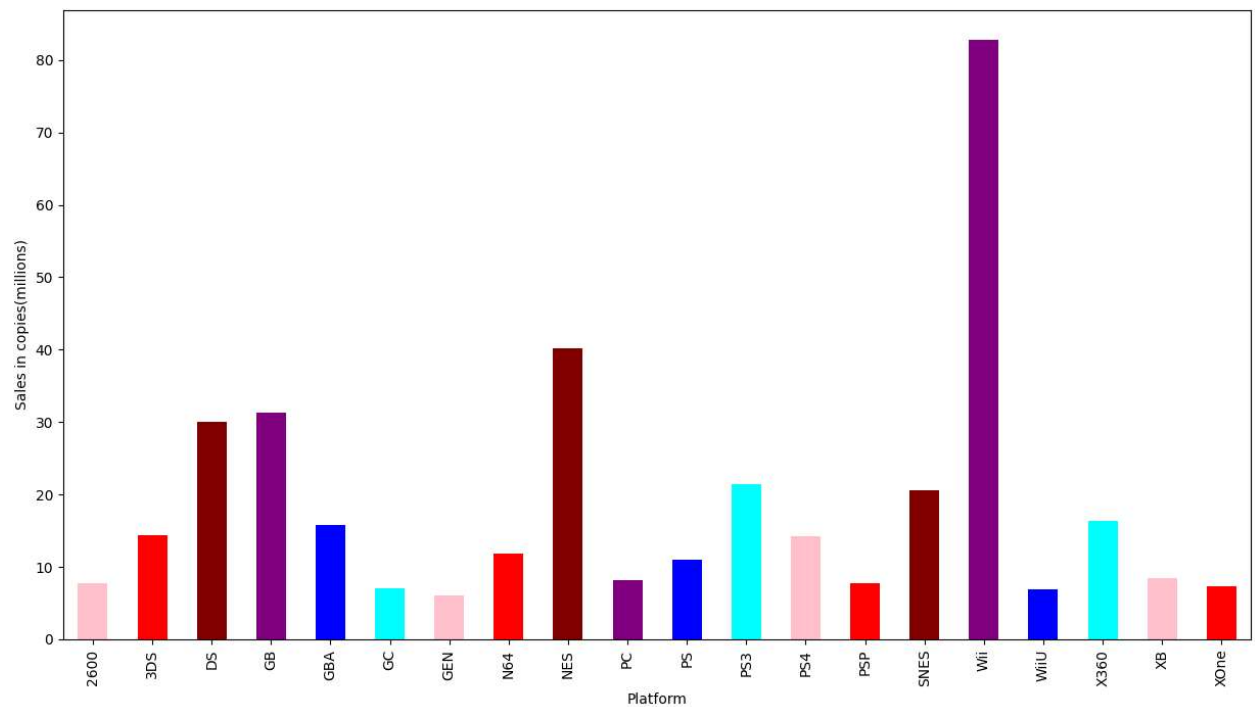


Nintendo have long been at the forefront of innovative video game development, when video games were already deemed to have been a passing fad in the mid-eighties. Nintendo brought out the NES, and it came with Super Mario Bros, Duck Hunt and the associated light gun accessory (1). This is reflected in the above two graphs at 1985. The biggest anomaly on this small plot of the top ten best sellers is Wii Sports another Nintendo invention, which sold 82 million copies, further examination reveals that the 2006 edition of Wii Sports came packaged with the Wii Console on its release in 2006 (2).

Such is the impact of the Wii in 2006 that it also skews the Genre popularity information for this period due to the huge volume of Wii Sports sold. As shown in this bar graph:

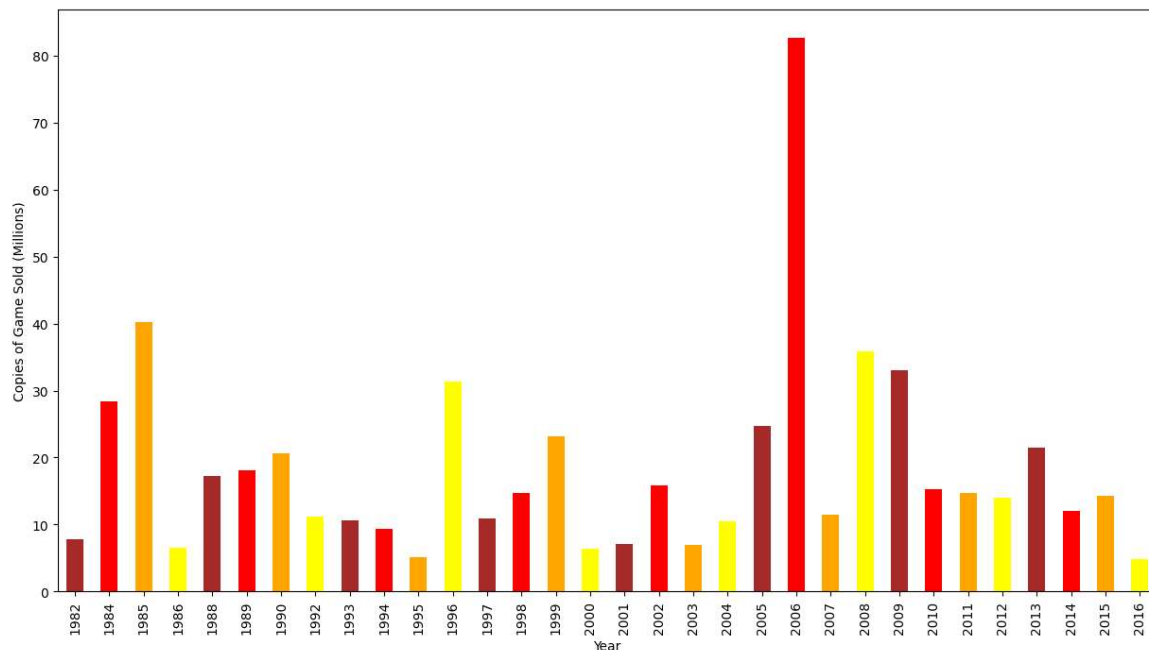


Shown in blue, the Sports column and to a lesser extent the Platform column (Super Mario Bros for example) clearly demonstrate a significant spike in sales, and to corroborate this interpretation of the data, we also see a spike in Nintendo gaming platforms over others on the following bar graph:

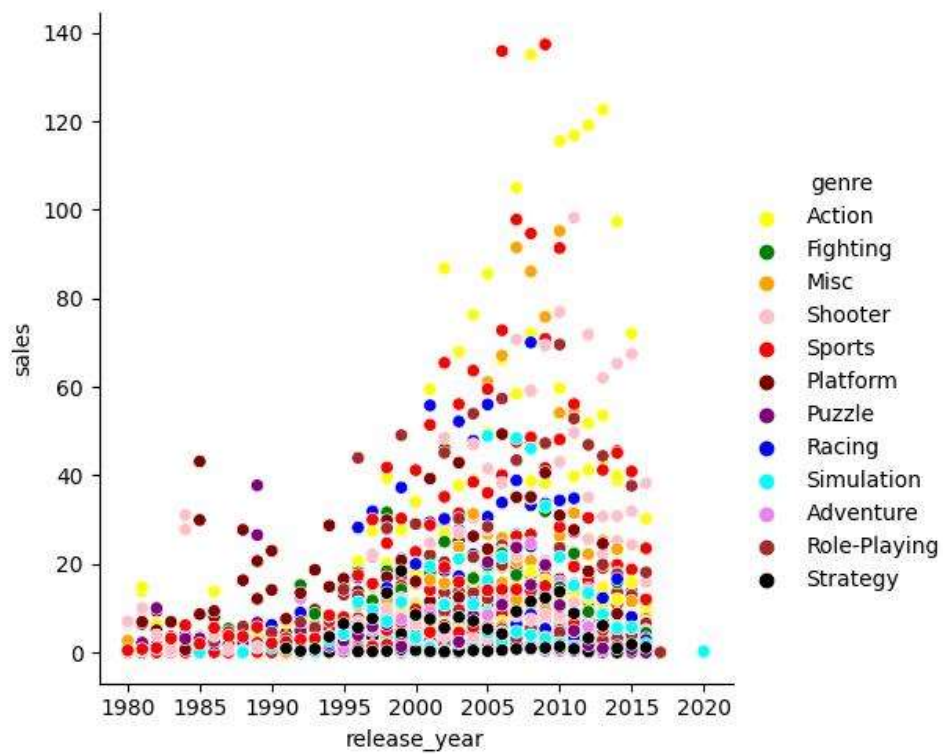


The popularity of the Wii as a gaming platform is clearly demonstrable in purple, while NES, DS and Gameboy (also Nintendo) platforms showed strong performance also. It is also worth mentioning that Tetris and Pokemon Red&Blue (from our ten bestselling games above) are both Gameboy games(Nintendo).

The same information can also be gathered by comparing the yearly sales where we see the same spikes.

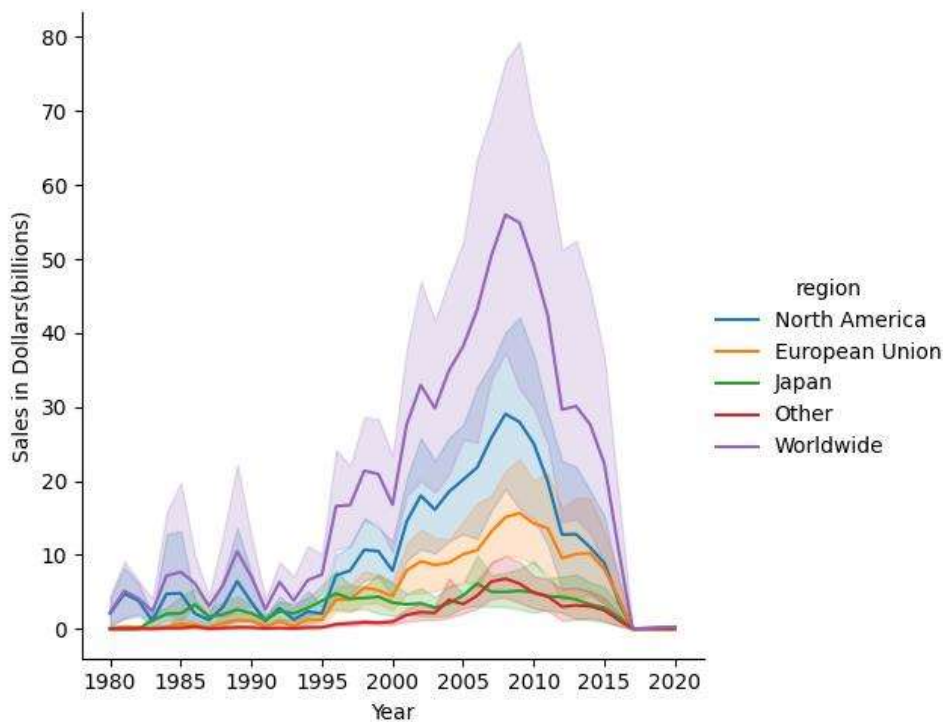


From this one Dataset, it's very difficult to see any evidence of high-ranking games outside of Nintendo's scope. This is why we are consulting other datasets to gather insights into gaming. In this dataset (Dataset: year_genre_region.csv), we can view a slightly broader view of genre sales by global region:

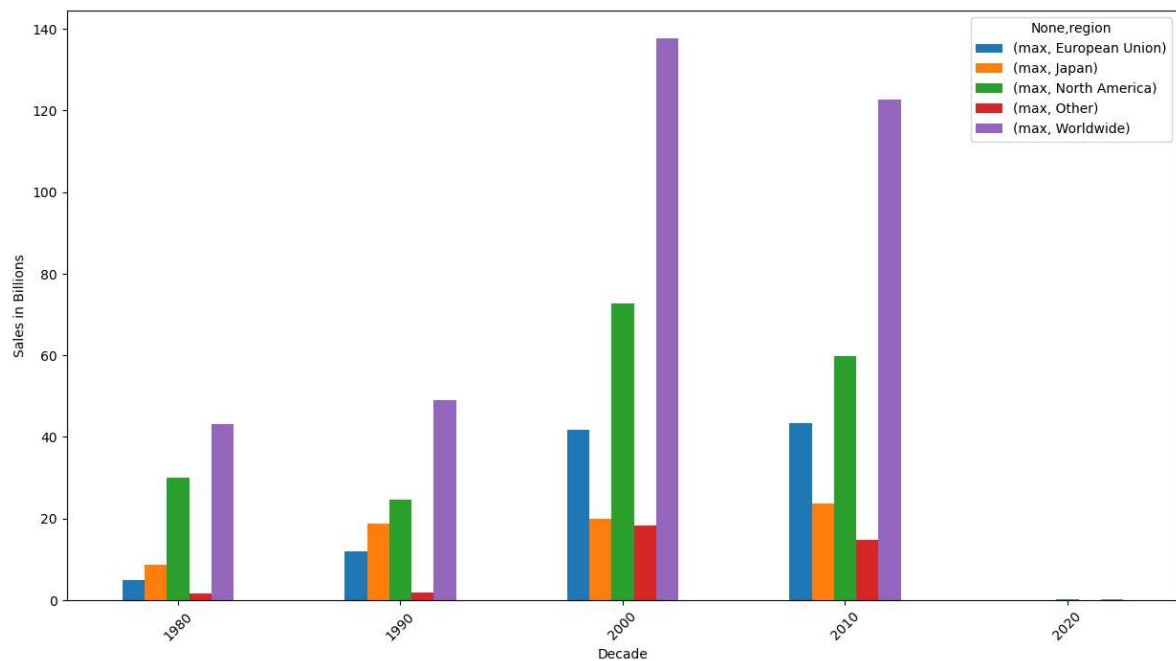


While Platform and Sports games still make a strong appearance, it is clear from this graph that Action, Shooter and even Misc games have shown strong sales in recent years particularly.

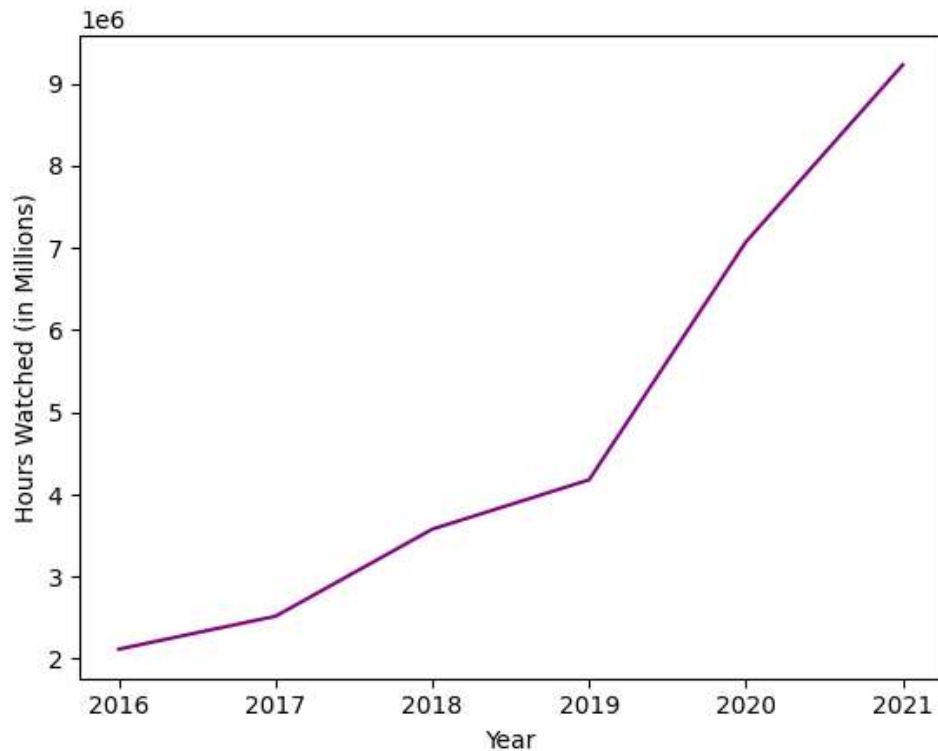
We can also see from this Dataset global and regional game sales:



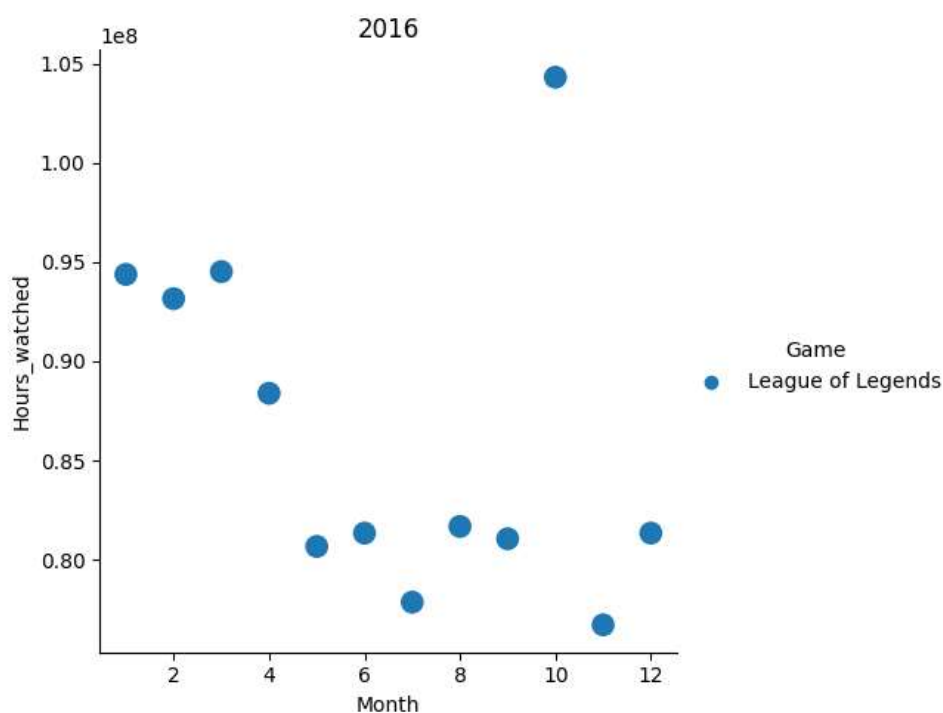
The above line graph seems to suggest a tapering off of sales in the late 2010's but it is my opinion that this is in fact due to incomplete data in those years, they have been left in so as to not compromise the integrity of the data. It is clear from the data that there has been steady growth in the game industry since its inception in the eighties. What is interesting to see also is that while North America has always been a large consumer of games, Europe has shown a distinct growth since the late nineties, possibly reflecting Europe's technical advancement but more likely reflecting the growth of the EU as more member states joined and increased its consumer power as a unit. The following bar chart shows Europe and North America's huge increase in game sales in the 2000s. It is interesting to note that Japan, despite being the manufacturer of many of the games and platforms we have looked at, shows modest sales growth in this dataset.



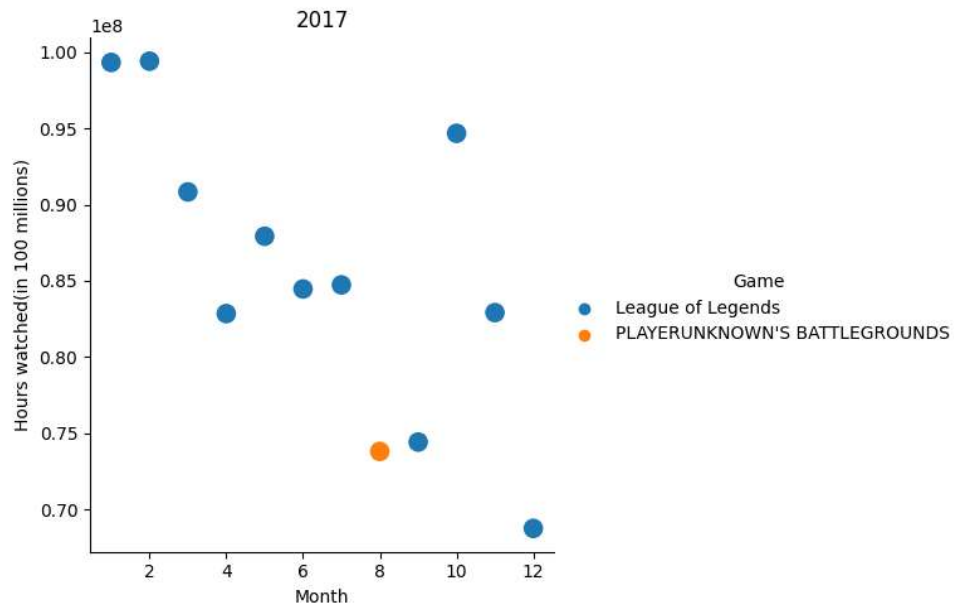
We will now move onto Streaming and specifically the impact of Twitch on gaming. From this line graph we can clearly see a trend. Between 2016 and 2021 there has been a steady increase in hours watched on Twitch. While Twitch is not specifically a gaming streaming service, it certainly has a gaming population in terms of streamers and viewers.



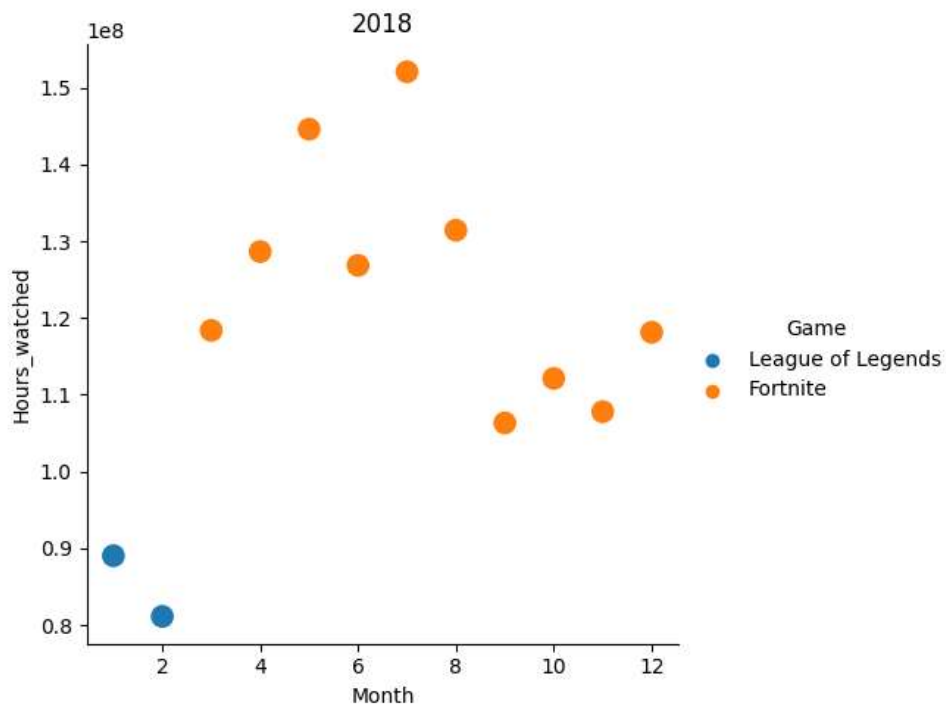
A point of special interest with Twitch is how often games that are Free to play, come out on top as the most watched. Since 2016 League of Legends has been a regular favorite on Twitch, at any given moment there are tens of thousands of people watching people stream their live gameplay. League of Legends is one of many games now that market themselves as 'Free to Play with in-game purchases.'



In 2016 (graph above), League of Legends was the most watched game on the platform every single month.

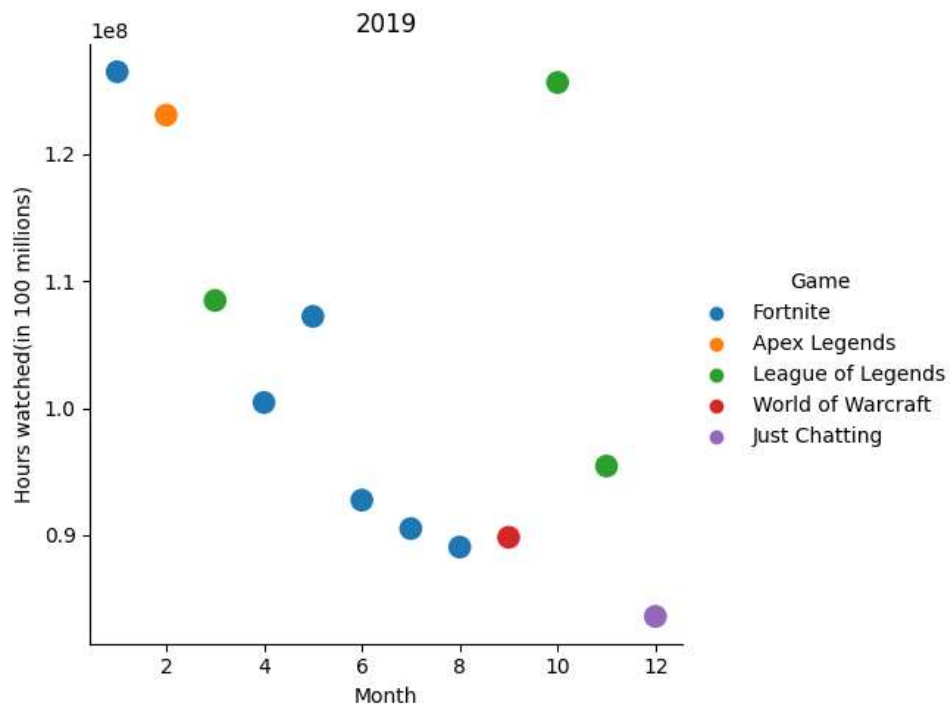


In 2017, League of Legends(LoL) only missed out on being top ranked game once, losing out to PlayerUnknown's Battleground (PUGB) when it was newly released (3). Interestingly, PUGB has now also moved to Free-To-Play. Moving onto 2018 we can see another success story of the Free-To-Play revolution:

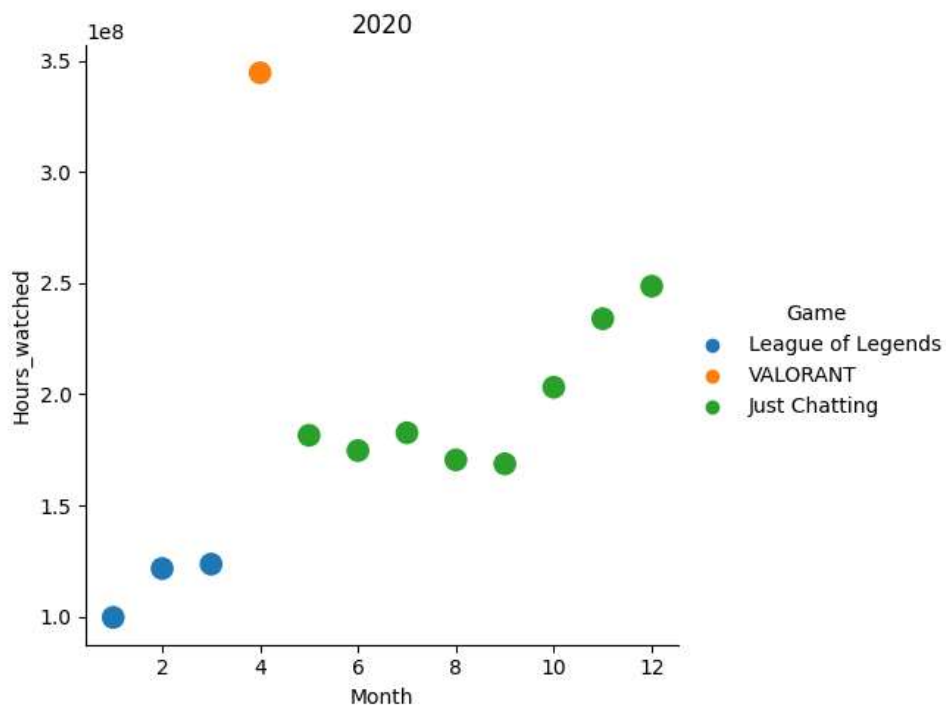


Fortnite was an early innovator as a game that is free to play but where you pay for Season Passes, Cosmetic modifiers and emotes for your characters. Fortnite is also classed as a Survivor game, its basic game-mode is a battle royale/ last man standing. This is basically an evolution from a first-person shooter game such as Call of Duty (which incidentally brought in a similar game mode after Fortnite's success). Survivor /Battle Royal was not a genre option in our original data source, the technology (and the high-speed internet connection requirement) didn't really exist before the 2000's and that is also one of the main reasons that live-streaming was not a viable (or even conceivable) option before this period.

What we see in the 2019 data is that Twitch is now wide-open, genre wise, any new game could claim top rank for a month but free-to play games are here to stay, all the games in the 2019 data are free-to-play some extent.



The 2020 data goes further in that it clearly shows the cultural significance of Twitch:



The year starts off normal enough with LoL taking the top spot January through March, April sees the release of the closed beta for Valorant. The rest of the year, due to a certain global event, is given over to Just Chatting. A self-explanatory title.

Conclusion and Insights

This report has endeavored to spot significant trends in how people buy, play, watch, and generally consume games. We have demonstrated how big companies with new platforms and controllers, as well conglomerate marketing budgets can skew the data. It is very difficult for small and/or independent game development companies to compete with a manufacturer who can also produce accompanying movies, TV shows, consoles and merchandise but sites like Twitch can make the market a bit more egalitarian to a degree. The popularity of PC gaming also means that is easier for gamer developers to put their games on the market through sites like Steam, instead of having to try and make a deal with big console manufactures like Nintendo, Sony and Microsoft. However, going forward it is not going to get any cheaper or easier to analyze gaming data. Consumers and companies are more protective of their data than ever before, streaming platforms like Twitch evolve and change all the time, and the bigger they get they more money they need to make so that is going to continue to change the way Twitch works, advertisement revenue and tiered subscription plans. Consumers will move away if it becomes too costly and it will be difficult to predict where the next big influence on gaming will come from.

In inspecting these datasets, I found that there was a combination of both numerical and categorical data, going forward with machine learning I would like to use Classification methods to work out upgoing trends in genres and compare prerecorded content like YouTube with Live Streaming as on Twitch. There are certain areas like competitions where live stream has an advantage. There's less list of embarrassment I presume with prerecorded content. They both have their advantages.

Regression machine learning methods could be useful to compare sales of optional addons in free-to-play games, to see whether the cost/benefit analysis of creating new content and season passes vs making cosmetic changes to existing characters/code.

Games are well on the way to becoming a respected form of media, with new genres evolving all the time (4). This report hasn't even touched on mobile games and the huge reach they have that makes them more inclusive than other platforms. This report has hopefully demonstrated the breadth and reach of gaming through data analytics.

References

- (1) <https://www.sofasandsectionals.com/history-of-nintendo-entertainment-system>
- (2) <https://www.britannica.com/topic/Nintendo-Wii>
- (3) https://store.steampowered.com/app/578080/PUBG_BATTLEGROUNDS/
- (4) <https://www.polygon.com/features/2019/11/8/20943571/new-game-genres-decade-in-review-2020-battle-royale-fortnite-dark-souls>