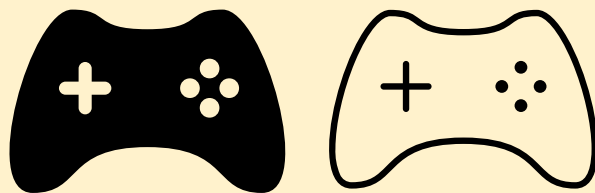




Billy Lacey  
December 2022  
Coding Temple - Capstone I

# Video Game Sales Analysis





### Data Notes:

- Credit goes to Gregory Smith who uploaded the .csv file I used for analysis to Kaggle. Data comes from a web scrape he did of vgchartz.com which had all the video game sales data. You can find dataset at <https://www.kaggle.com/datasets/gregorut/videogamesales>
- Note all sales figures are denoted in millions (ex. 40 would represent 40 million in sales)
- Sales figures are as recent as 2016, but it was partial year in terms of revenue recorded, 2015 is last full year of sales data
- If video game title released on multiple platforms then there are multiple records for a single game in .csv

I thought this was an interesting data set when I came across it on Kaggle. I looked for something that was more recent than 2016 but was unable to do so. Looks like another web scrape of vgchartz.com would do the trick, so will probably circle back after program is complete to see how the Kaggle user performed the first scrape and give it a try myself. With that being said with 35 years of sales data from the top regions in the world, I'm sure there are plenty of valuable insights to gain. By no means am I a gaming expert but video games were a staple in my childhood growing up. I first had a PS One and was a sucker for the Spyro games, ever since then I stuck with PlayStation consoles but through friends I've played on numerous different consoles (handheld and home) so I have good understanding of the context surrounding the data but I'm sure there are going to be many things that surprise me along the way.

### HYPOTHESIS:

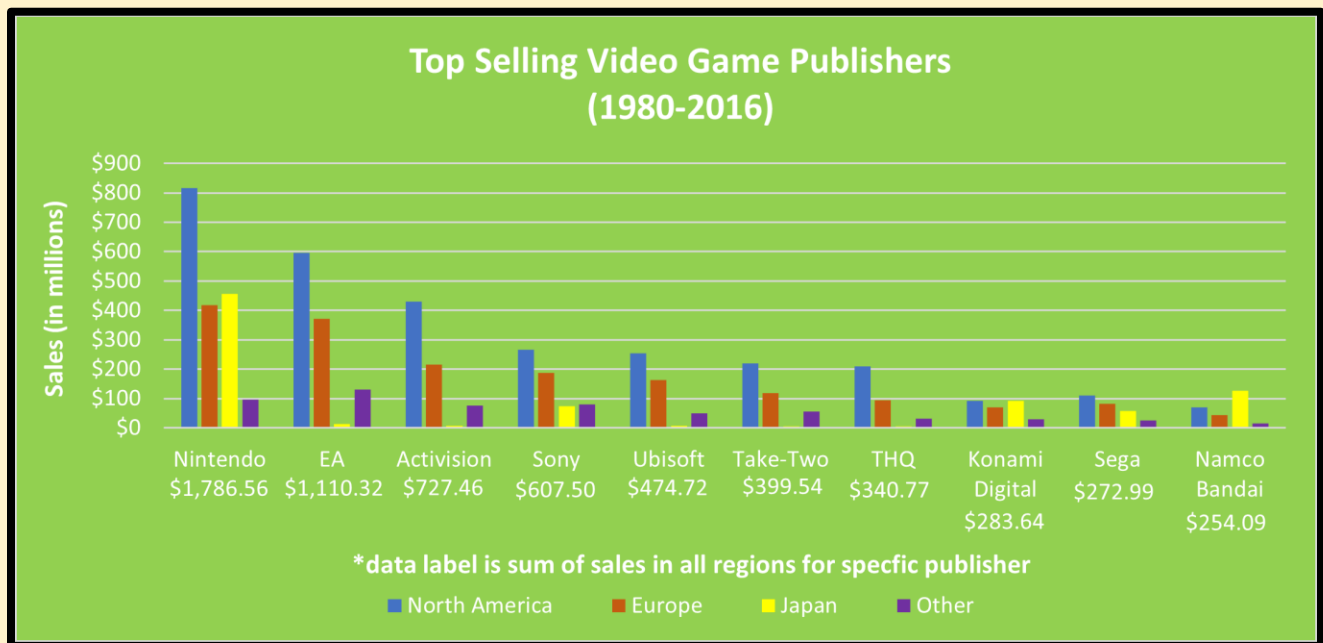
- Nintendo will be top publisher in Asian/Pacific regions and most likely will be top publisher globally
- PlayStation 4 (PS4) will have the most sales among all the different platforms (**note: this is not unit sales of individual consoles, but amount of sales revenue generated from games released on that platform**)
- Sports games are going to be the top selling genre (**note: racing games fall under my sports genre umbrella**)
- Using 1990 annual sales as base and comparing to 2016 annual sales I think you'll see an industry growth of 4x-5x in terms of game sales



- Out of top regions in world, North America will biggest market just because of population size, but I wouldn't be surprised if Japan region was bigger since gaming culture there is much stronger
- For each major platform umbrella (handheld and home) I think the top franchises in terms of total sales will be as follows:
  - Nintendo: Mario Kart games
  - PlayStation: Grand Theft Auto Games
  - Xbox: FIFA games
  - PC: World of Warcraft games

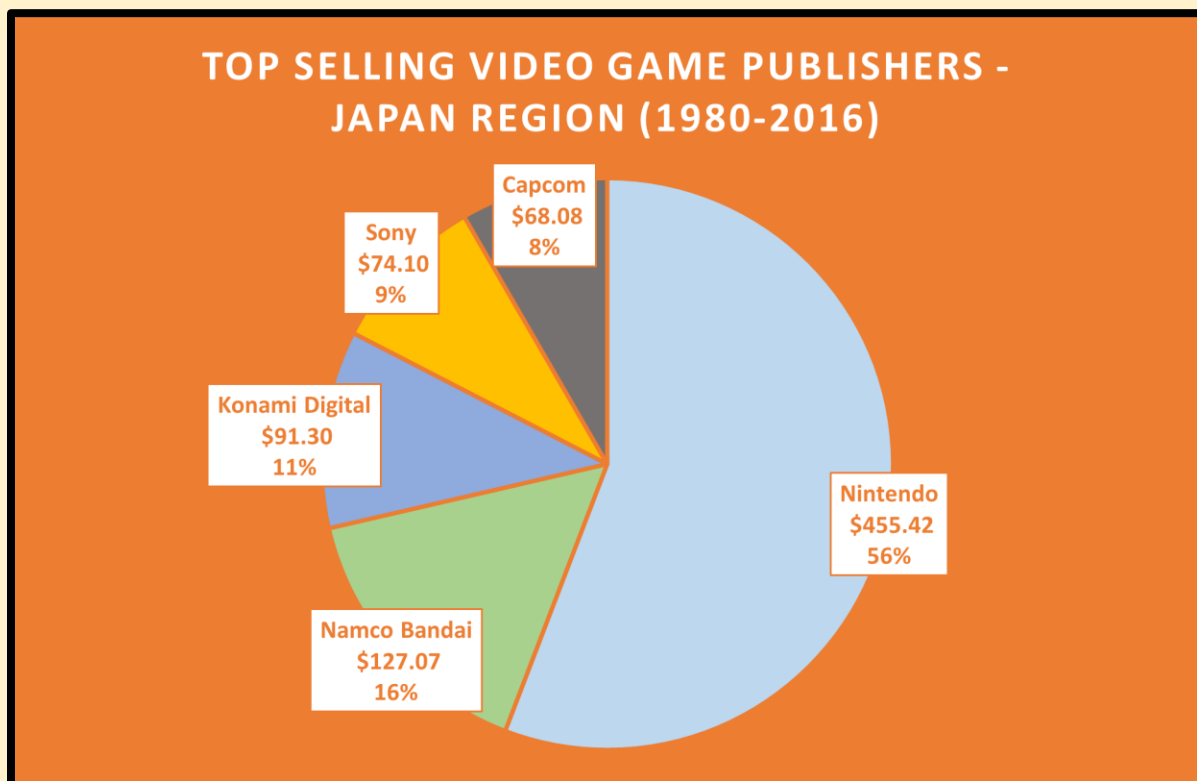
## I. Top Selling Publisher

Well Nintendo certainly seems to know what they're doing, they are by far the top grossing video game publisher with lifetime game sales of roughly \$1.8 billion. The only other publisher to have over a billion in game sales was Electronic Arts, which earned roughly \$1.1 billion in revenue. To give you sense of dominance Nintendo has in this category, their lifetime video game sales revenue is essentially equal to that of the remaining publishers in the top five of game sales. Take a look below:





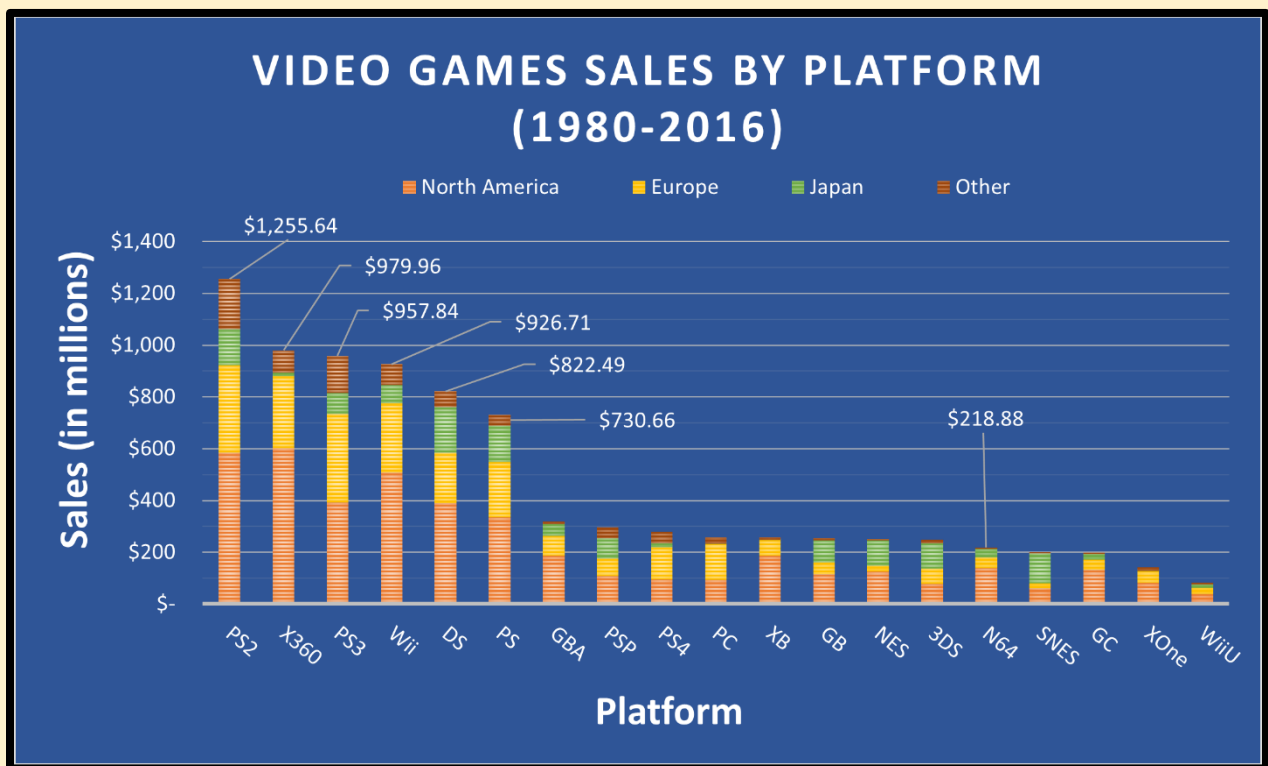
As you can see Activision, Sony, and Ubisoft (the 3<sup>rd</sup>, 4<sup>th</sup>, and 5<sup>th</sup> top selling publishers) combined global game sales is \$1.809 billion where Nintendo individually grossed \$1.786 billion in game sales. The graph above also breaks down which regions were responsible for a publisher's game sales. One thing that must be taken into account is that Nintendo was a pioneer in the gaming industry, they released their first game in 1983 on their NES system while others were a tad later to join the party. Other than Activision (which I was surprised to learn released their first game before Nintendo in 1980), the other top publishers didn't release their first game until a decade after Nintendo. Electronic Arts (EA) released their first game in 1992, Sony didn't release their first game and flagship home console (PS One) until 1994, and Ubisoft was the latest out of the bunch with their first game release in 1995. I only bring this to attention to help explain the revenue discrepancies among the top publishers. Another aspect of my first hypothesis was that Nintendo would dominate in Japan region and it turns out that was the case. Out of the top 5 publishers in Japan, Nintendo had over a 50% market share in game sales, with lifetime game sales of \$455 million the next closest publisher was Namco Bandai who had a 16% market share and \$127 million in game sales. More info below:





## II. Top Selling Platform

Well my hypothesis about the PS4 being the top performing platform in terms of game sales was comically off. I should've taken into consideration that this dataset only had sales data up to 2016 and the PS4 wasn't released until the winter of 2013/2014; with only 2 full years of game sales it definitely was not the top platform! I did think it was surprising that it had already broken into the top 10 though in those two years, with game sales at \$278 million it was easily beating the Xbox One, which released at a similar time and only had \$141 million in game sales up to 2016. Sony did have the top selling platform it just wasn't the PS4, it was the beloved PS2 in all it's glory. The PS2 had \$1.25 billion in game sales, the only platform to achieve the feat of breaking \$1 billion in game sales. The graph below gives you better idea of how different platforms performed:

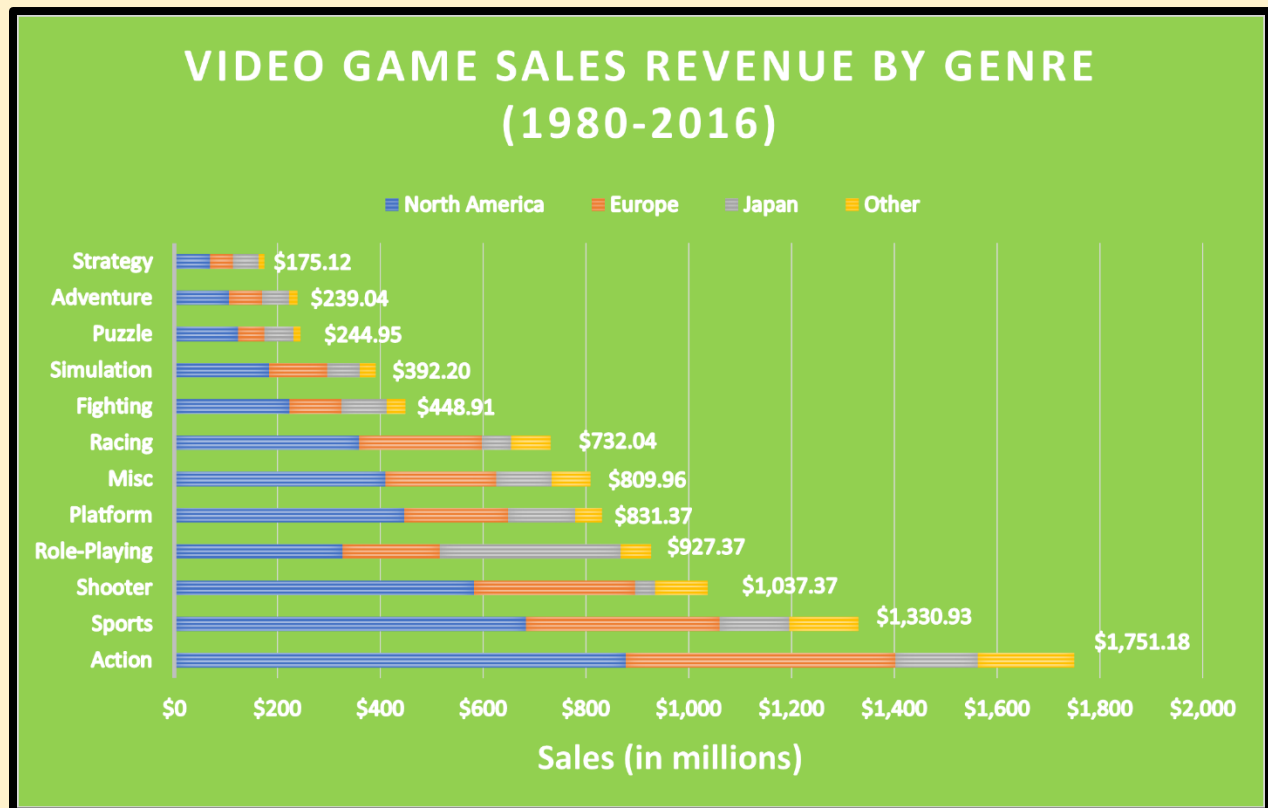


**Quick Notes:** Sony had 5 of their 6 platforms in the top 10  
Nintendo had 9 platforms to gross over \$200 million



### III. Top Selling Genre

Is Racing a separate genre or a sub-genre of sports? It may seem pretty trivial but depending on how you classify those games it changes the result of the top selling genre in game sales, see below:

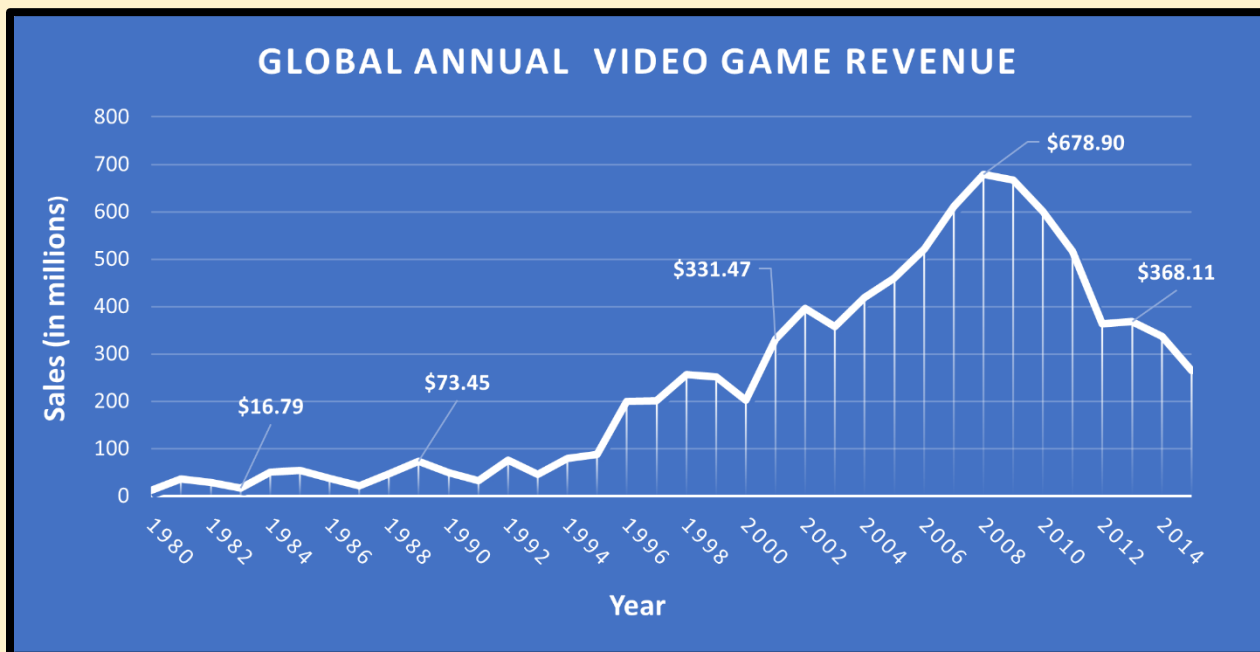


As shown in the graph, Action games have grossed roughly \$1.75 billion in lifetime sales. The most out of any genre with a sizeable gap between Sports (the next closest genre) which has roughly \$1.33 billion in game sales. You will notice if you include Racing game sales (\$732 million) with Sports their collective total is over \$2 billion in sales though, how you want to determine which genre sells better is up to you, it ends up being subjective with what games you think belong grouped together. One could argue Shooter games belong under Action games so forth and so on. I think its best to leave the sub-genres classified the way they are and classify Action as the best selling genre. One thing about the Shooter genre I found interesting; it is the 3<sup>rd</sup> best selling genre globally but in Japan region it is the lowest performing genre of ALL the genres. The reasons behind that stat would require more analysis and of data sets that I don't have at the moment, so I won't speculate as to why but I thought that bit was the most interesting when looking into the different genres.



## IV. Industry Growth

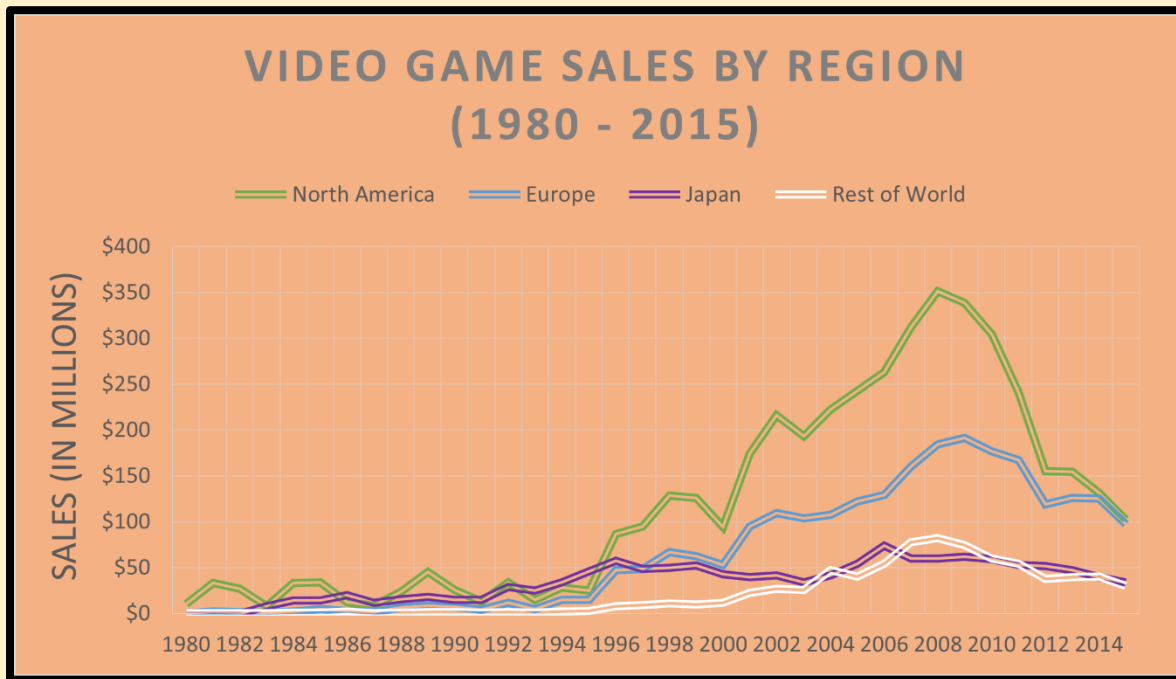
Looking at annual video game sales throughout the years was a really interesting query to dive into. One could assume that with the rise of the popularity of video games over the past 4 decades that game sales volume would also increase. This trend was exhibited in the data but starting around 2008 global annual game sales started dropping. This stuck out to me because I know to this day the video game industry is still growing, it's one of the largest industries in the world, so why wasn't I seeing that correlation in the data. My guess is that this is when you start to see the rise of digital game purchases and or some other new type of sale that isn't being taken into consideration. In addition, when conducting my initial query I needed to adjust my parameters to get relevant results back. I started by looking at 1985 to 2016 but 2016 wasn't a complete year of sales in the .csv file and 1985 didn't have a large enough sample size in my opinion (it was about 30 games). I ended up comparing the years 1995 (\$88.11 million) and 2015 (\$264.44 million), there was almost an exact 300% growth in game sales volume. Take a look at the annual video game sales volume throughout the years:



You can see in the graph the steep decline in game sales I was referring to above. The top selling year for video games was 2008 which had \$678.90 million in sales. Using that sales figure compared to 1995, the size of the industry grew 770%, vastly different



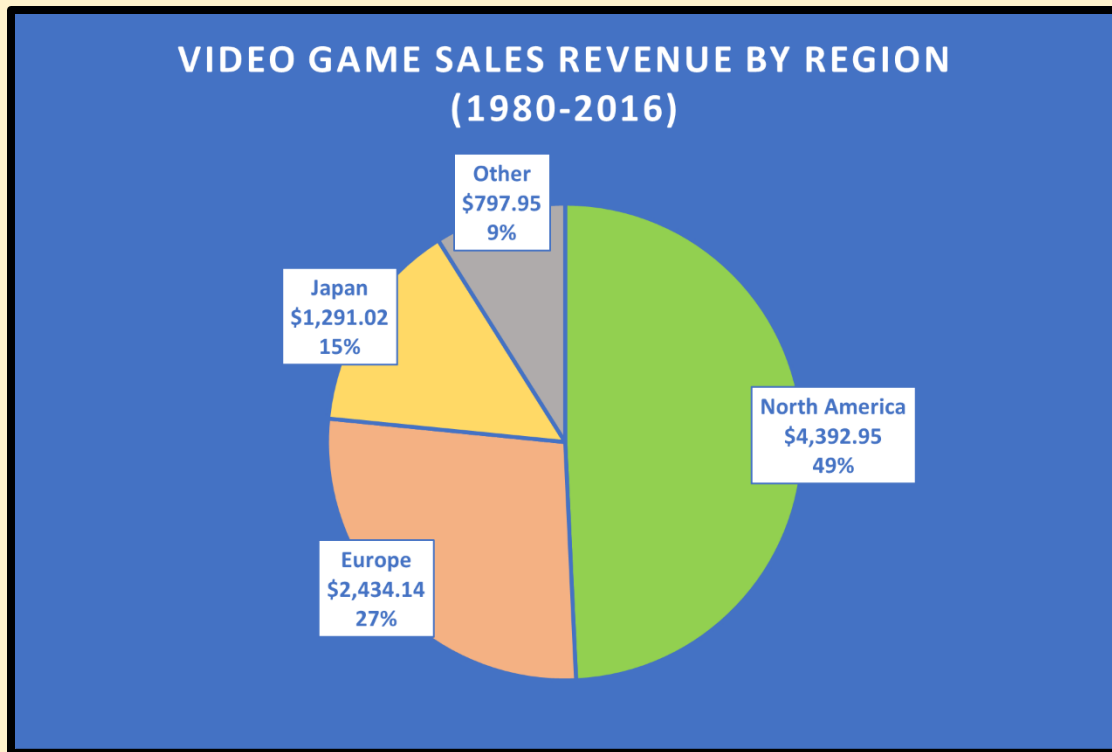
then 2015. That's an area that needs to be explored more but I would need another dataset. I also thought it was interesting to see how different regions performed in game sales throughout the years:



## V. Top Selling Region

I think of all my queries this was the least interesting, although I was a tad surprised at how small the Japan market was relative to North America and Europe. North America has the most sales of all regions heavily due to the fact that it is the largest region. North America is responsible for roughly \$4.4 billion in total game sales, that's a 49% share of total revenue generated from game sales. Europe was next region with \$2.4 billion and Japan was third with \$1.3 billion in game sales. You can see the dominance North America has on video game sales in pie chart below:



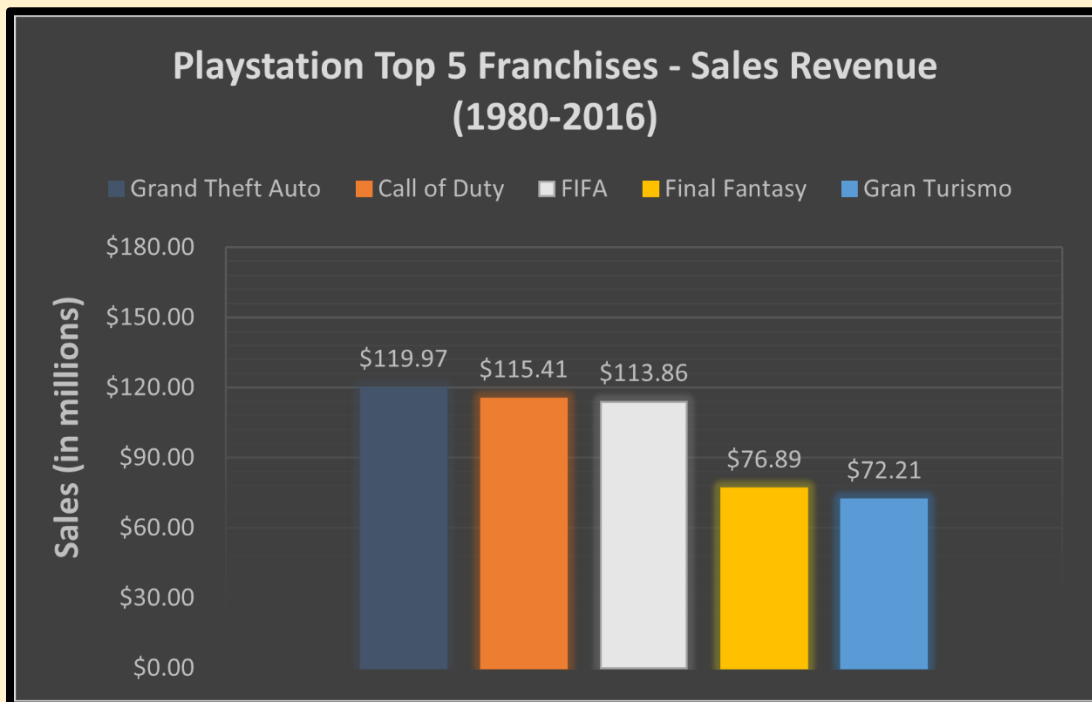
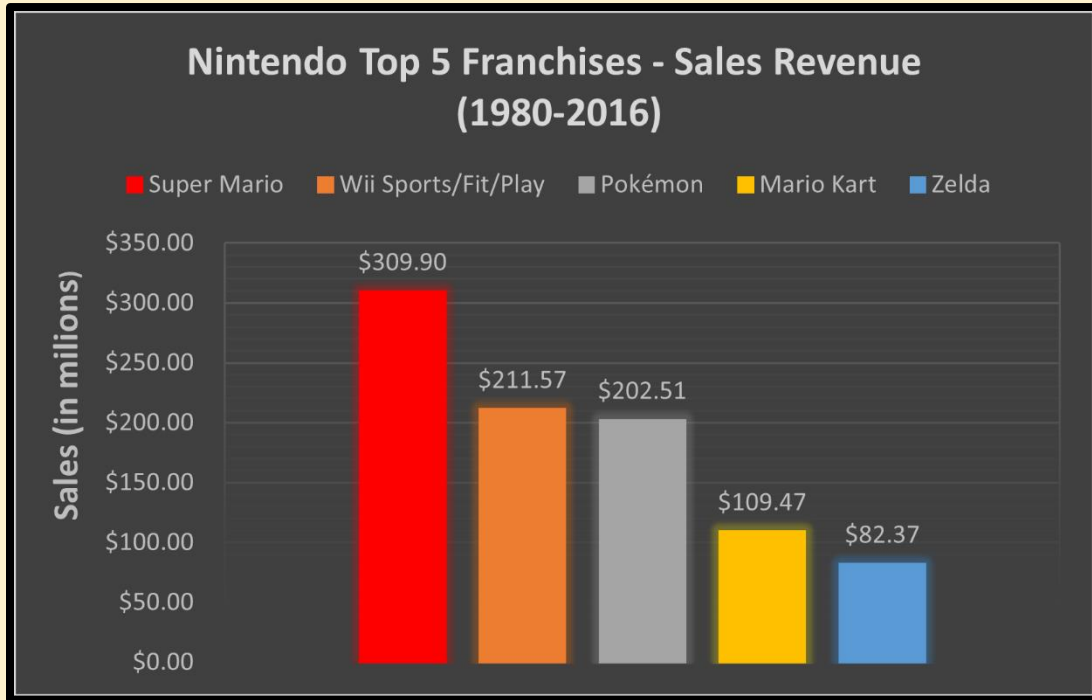


## VI. Top Selling Franchises

This was a fun query to dive into, because I was looking at a lot of different games that took me back to when I was younger and was also surprised at how little some of them grossed compared to other franchises even though they may be praised critically (ex. God of War, Fallout, etc.). I made four different graphs for each major platform umbrella; those being Nintendo, PlayStation, Xbox, and PC. The way I found these results is by first looking at top 30-40 selling games for each platform and then using `contains()` in Python and `grepl()` in R to group different title variations. After grouping by franchise I then summed global sales for each game to get the sales figure you see in the graphs. I also examined the franchise catalogs to make sure games were not getting counted twice (ex. Wii Mario Kart showed up in Wii Sports catalog but also Mario Kart catalog, it's more associated with Mario Kart franchise than Wii Sports/Fit/Play so I removed it so sales figures weren't counted twice). It was also interesting to me how much PlayStation outsold Xbox in similar games, Call of Duty being the exception to the rule, a lot of titles that were released on both platforms



performed much better in sales on PlayStation consoles. Take a look at the different platform's top franchises below:

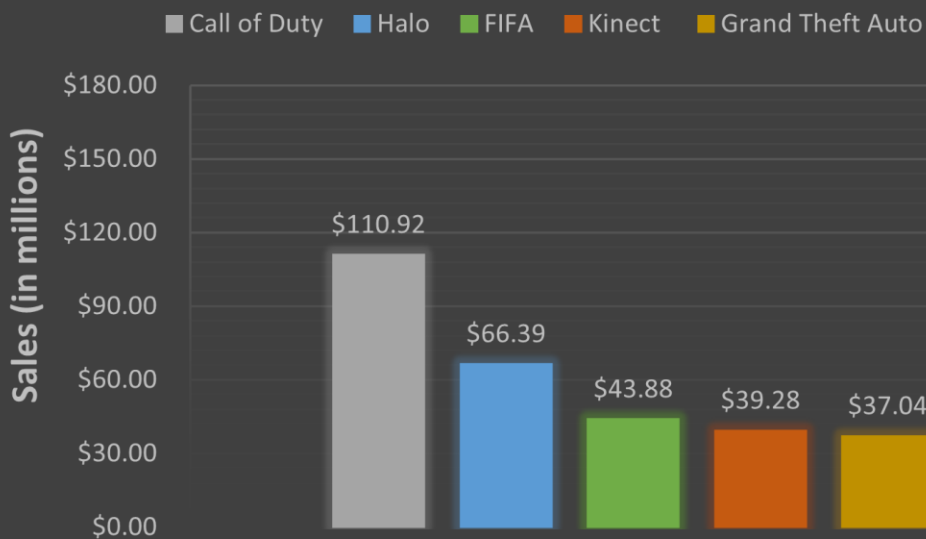




## PC Top 5 Franchises - Sales Revenue (1980-2016)



## Xbox Top 5 Franchises - Sales Revenue (1980-2016)





### Closing notes:

- *All queries were conducted with Python or R, mainly using pandas library in Python and dplyr and tidyverse libraries in R.*
- *Graphs were made in Excel by exporting a .csv of dataframe from Jupyter notebook.*
- *All files are included in github repo*