Google Colab:
https://colab.research.google.com/drive/1MK5MNn5XiTCbpzNtYEFiBCUnMMCe-Sd-?usp=sharing
Youtube:
https://youtu.be/Dc95P-peUJ0
Data taken from Kaggle:
https://www.kaggle.com/datasets/gregorut/videogamesales?resource=download

About the Data

This data contains a list of video games with sales greater than 100,000 copies sold. The number of sales is calculated in millions. Publishing dates go from 1989-2020. The data set consists of a mix of strings and floats. Less than 3% of the data set contains null values, which is mainly seen in the 'Year' and 'Publisher' columns. The data set has already been ranked based on sales and is in descending order. There are over 16,500 games listed in this dataset.

Assumptions consist of viewing the data set as a list of video games sold based on physical copies only. This will not include data from any downloadable version of these games. The focus, in terms of sales, will be averaged from the Global_Sales column. This gives you a general idea of worldwide sales which can easily be broken down and investigated further if analysis requires it.

|	Rank	Name	Platform	Year	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales
0	1	Wii Sports	Wii	2006.0	Sports	Nintendo	41.49	29.02	3.77	8.46	82.74
1	2	Super Mario Bros.	NES	1985.0	Platform	Nintendo	29.08	3.58	6.81	0.77	40.24
2	3	Mario Kart Wii	Wii	2008.0	Racing	Nintendo	15.85	12.88	3.79	3.31	35.82
3	4	Wii Sports Resort	Wii	2009.0	Sports	Nintendo	15.75	11.01	3.28	2.96	33.00
4	5	Pokemon Red/Pokemon Blue	GB	1996.0	Role-Playing	Nintendo	11.27	8.89	10.22	1.00	31.37
...	...	...	...	...	...	...	...	...	...	...	...
16593	16596	Woody Woodpecker in Crazy Castle 5	GBA	2002.0	Platform	Kemco	0.01	0.00	0.00	0.00	0.01
16594	16597	Men in Black II: Alien Escape	GC	2003.0	Shooter	Infogrames	0.01	0.00	0.00	0.00	0.01
16595	16598	SCORE International Baja 1000: The Official Game	PS2	2008.0	Racing	Activision	0.00	0.00	0.00	0.00	0.01
16596	16599	Know How 2	DS	2010.0	Puzzle	7G//AMES	0.00	0.01	0.00	0.00	0.01
16597	16600	Spirits & Spells	GBA	2003.0	Platform	Wanadoo	0.01	0.00	0.00	0.00	0.01

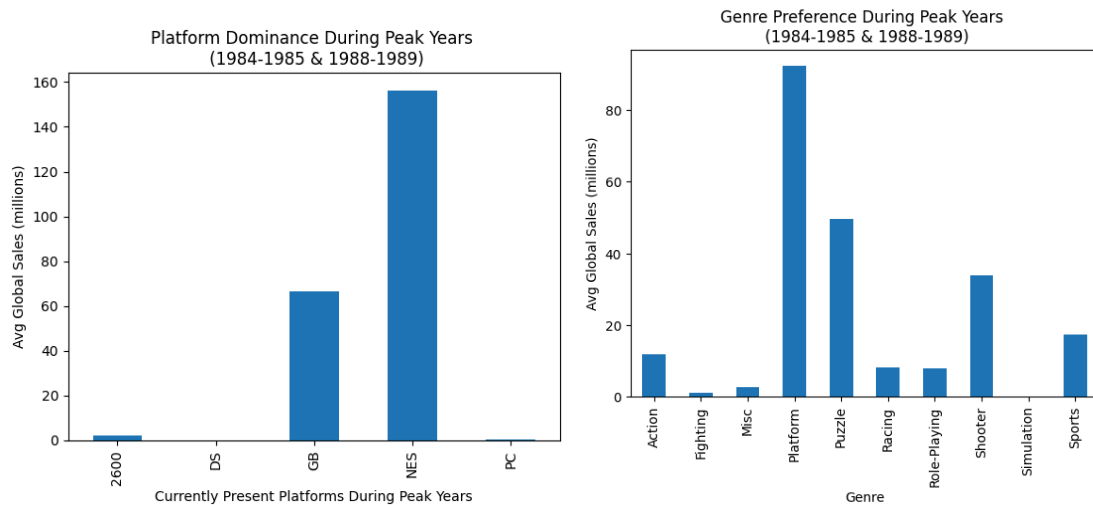16598 rows × 11 columns

Cleaning Data

Encoding Categories

To determine relationships between all variables except for the Name column, I needed to encode categories to make it numerical to put into a correlation matrix.

EDA

Avg Global Sales Throughout the Years
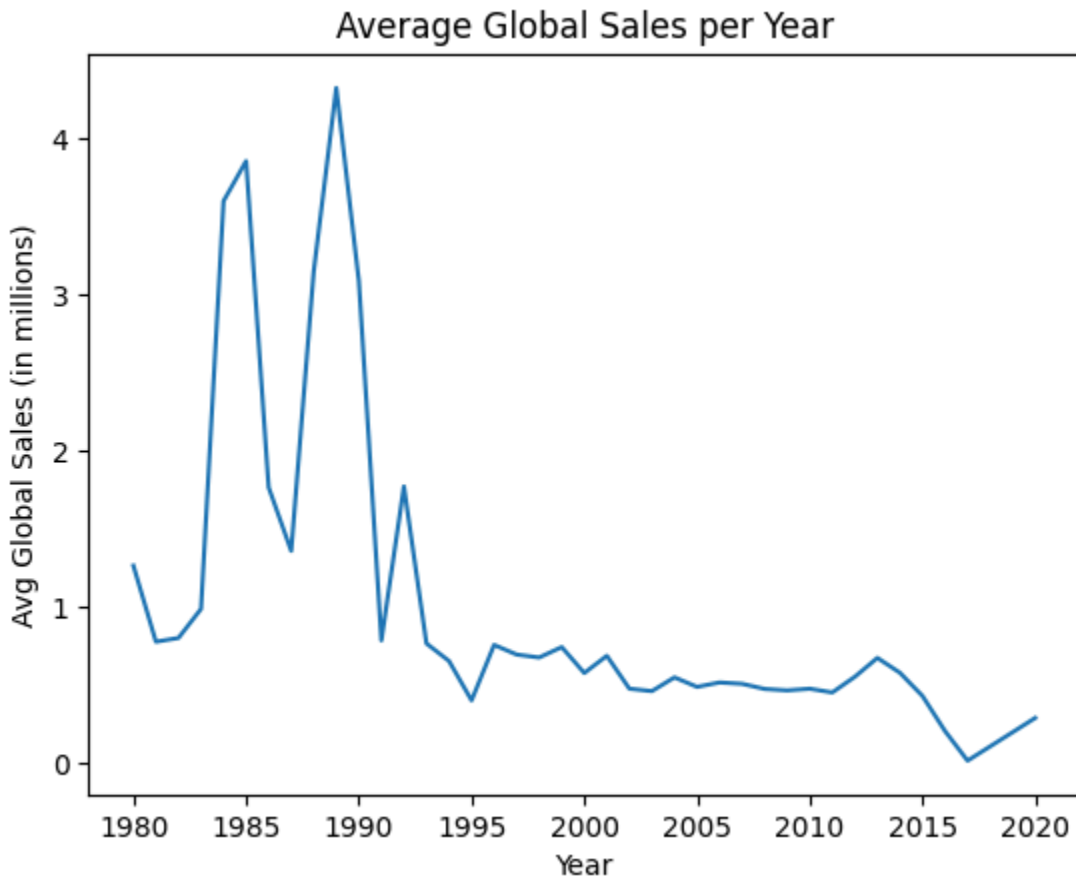
The average global sales per year show there was a huge spike in video game sales from 1984 to 1985 and then dropped drastically until peaking again at the highest in 1989. This could be because video games switched from arcade to console and many new companies were emerging and entering this market. There was a small peak afterwards but it ultimately never got as high as it did again. It continues a steady line from 1996 to around 2012-2013. There was an all-time low in copies sold around 2016. This could be a result of gamers turning away from physical copies and entering an era of online gaming and downloads.

Platform Dominance During Peak Years
(1984-1985 & 1988-1989)

Genre Preference During Peak Years
(1984-1985 & 1988-1989)

5 platform games were available during this period. Nintendo Entertainment Systems (NES) dominated the market during this time with around 156 million copies sold with the most popular games that we recognize to this day like Super Mario Bros with an average global sale of 40 million copies. Following Nintendo, GameBoy sales were a little less than half of NES' sales. Though the system isn't played with much these days, its games during the height of its era are still classics such as Tetris with an average of 30 million copies sold.

Considering the time and technology during this period, it makes sense that Platform games are the most popular games being sold. Puzzle, shooting, sports, and action games are included in the top 5 genres during the peak of copies sold. A platformer is a video game in which the gameplay revolves heavily around players controlling a character who runs and jumps onto platforms, runs over obstacles, and avoids enemies. This is usually depicted on a scrolling game screen. Think Super Mario Bros.

Global Sales in Peak Years
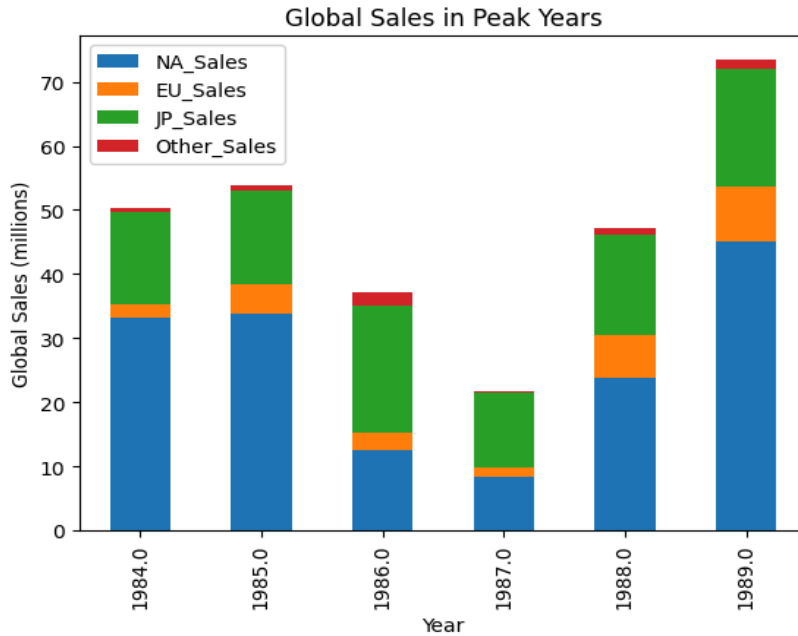
Overall, North America makesup for most of the global sales, and Japan comes in second. In general, it seems there was a jump in sales all over the world when publishers released some of the most popular games we play to this day.

Game Trends

Trend of Genre Preferences Over Time
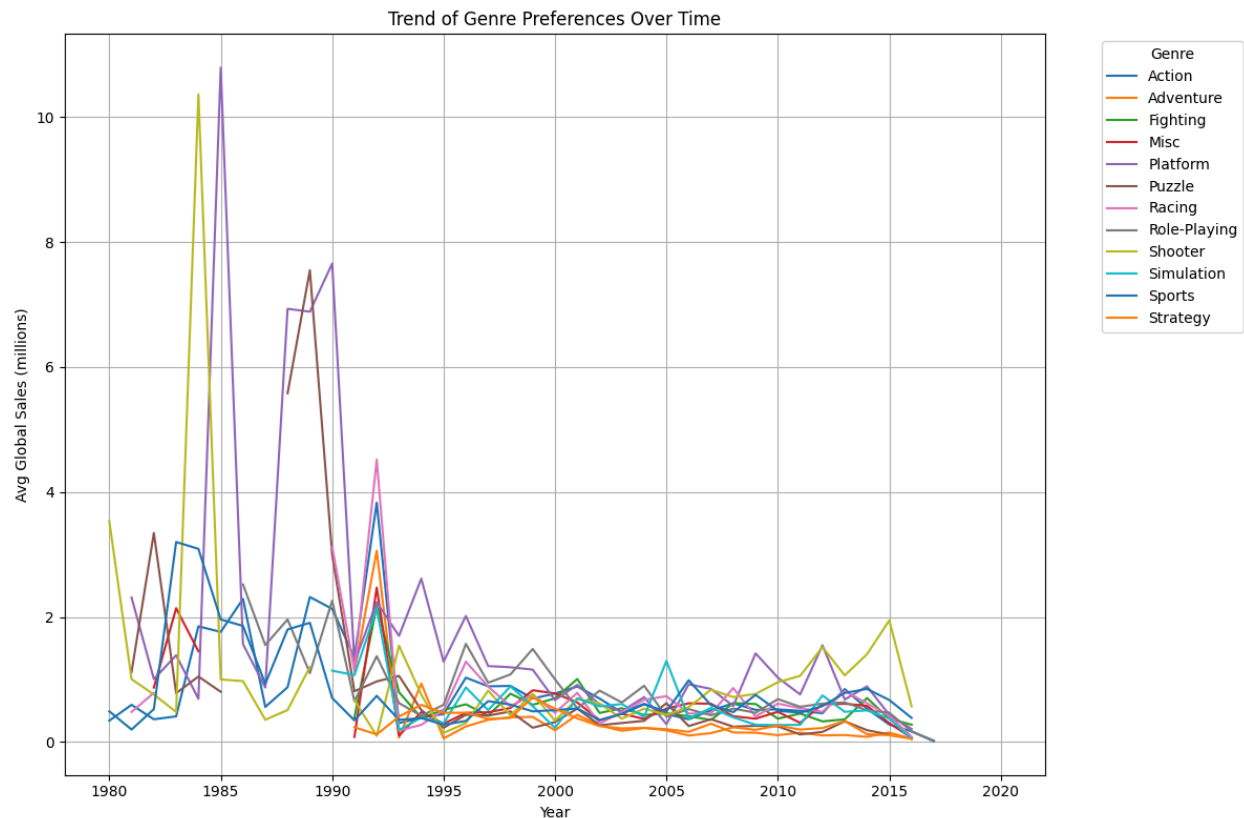
As technology advances and newer game systems come out, our genre preferences and platform popularity have changed over the years. For genre preferences, the top three genres that are still prevalent in today's games are Platform, Shooting, and Puzzles. As for platform dominance, the newer, the better. Keeping up with the Joneses is what it seems like when a new console drops. It started with 2600 to the Gameboy and Nintendo Entertainment Systems. In the 2000s, Microsoft came out with Xbox and Sony with PlayStation. These two game consoles are arguably the top two platforms to this day.

Platform Dominance Over Time
(Top Ten)