

Angela Vaynshteyn
September 8, 2019

Assignment 1 Video Games Sales Analysis

1. How many instances are in the data?

There are 132,776 instances in the data (132,784 including the header for each column)

2. How many attributes?

There are 8 attributes (Name, Platform, Year, Genre, Publisher, NA_Sales, EU_Sales, JP_Sales).

3. Are the attributes numerical or categorical?

There are four categorical attributes (Name, Platform, Genre, and Publisher) and four numerical attributes (Year, NA_Sales, EU_Sales, and JP_Sales)

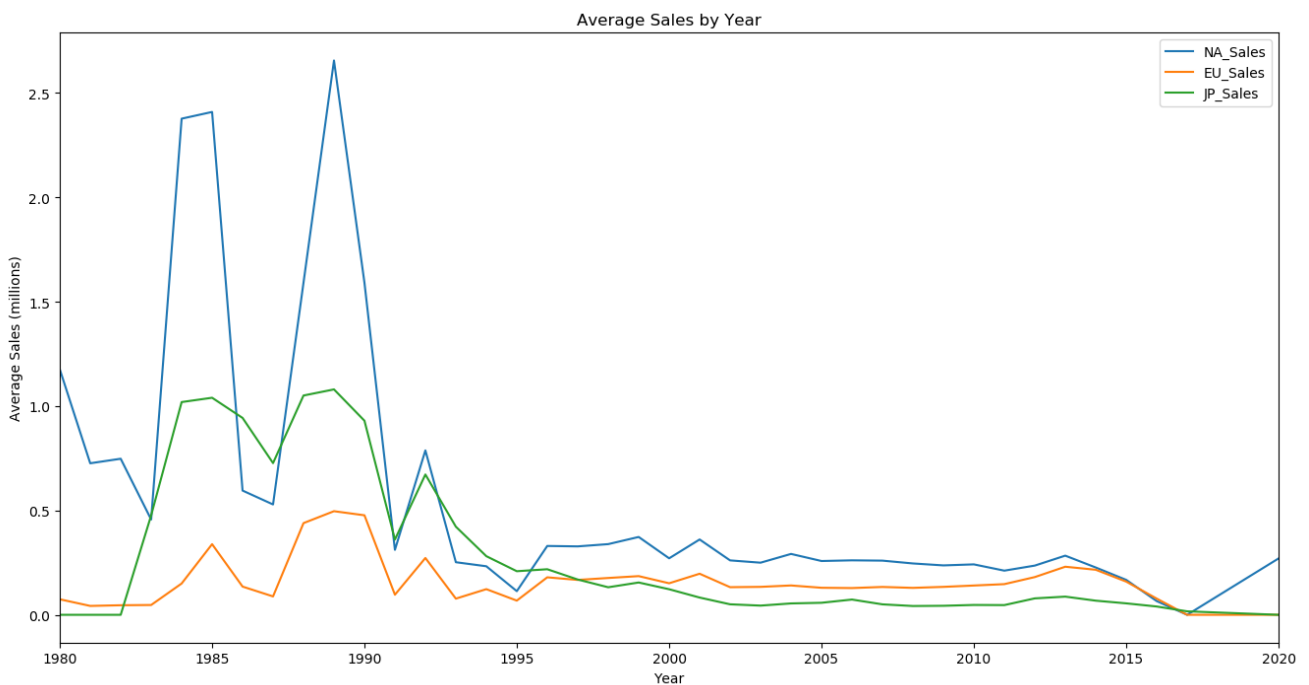
4. Any obvious outliers?

There are no outliers based on the code in rows 27-48.

5. Comment on any interesting trends you discover (describe the plots)

6. Include at least 5 different plots

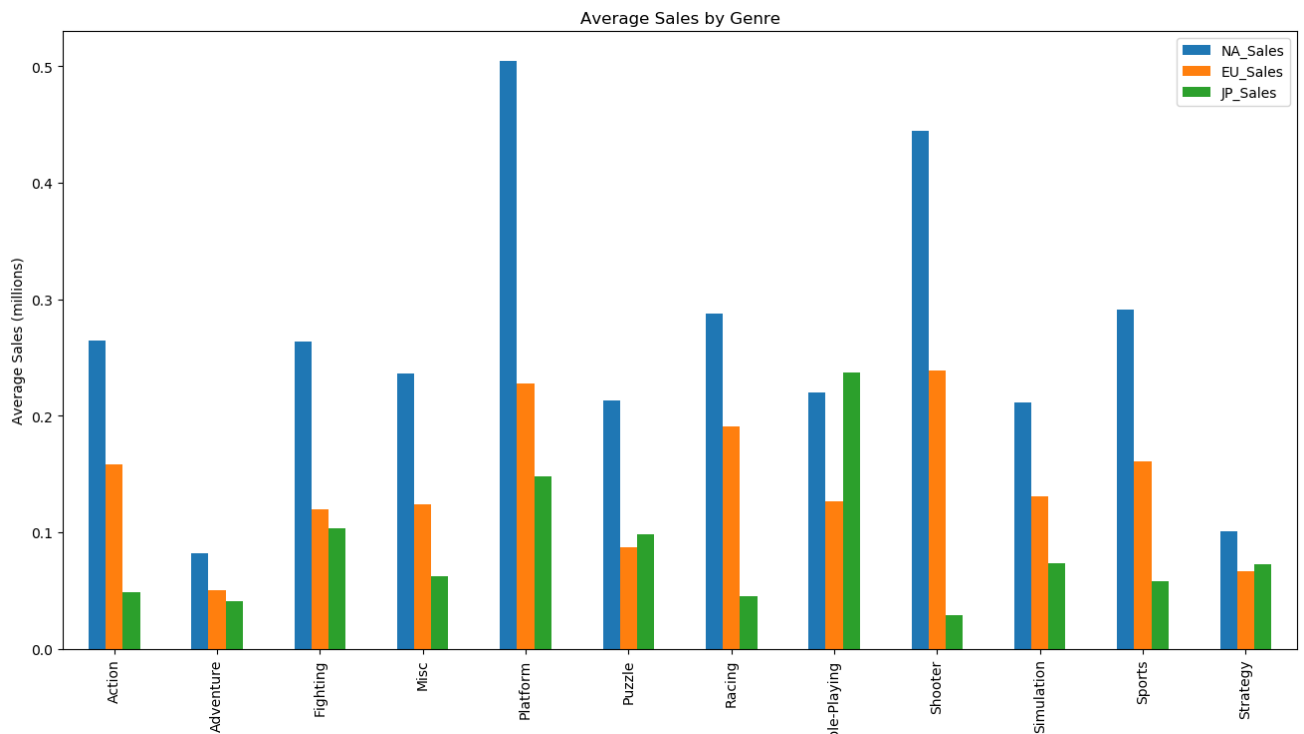
Plot 5.1) Average sales by year



The above is a graph for average sales (in millions) per year from 1980 to 2020 by every 5 years. Consistently, North America is the top consumer of video games. There was only a brief period after 1985 but before 1990 and another brief period right before 1995 when Japan spent more on video games. Japan remained above Europe at about mid-1980s to late 1990s. After 1995, Europe became a bigger consumer of video games than Japan. Before 1985, the sales of North America went down while the sales in Japan went up. There was a spike in the average sales per year in the US, Japan, and Europe right before 1985 and during 1990. It is possible that popular games such as Pokémon came out during those years, but that is not verified by

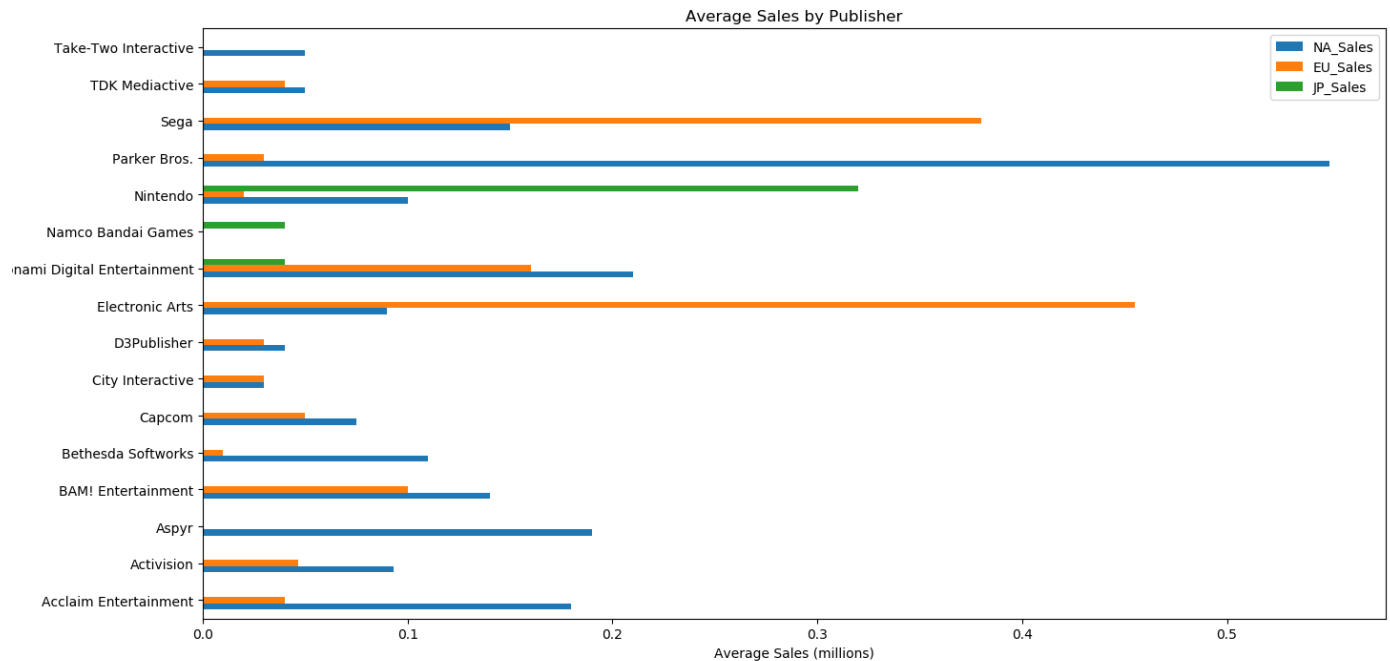
this chart. The sales in all three locations leveled out after 1995 and have remained the same since. Between 2015 and 2020, sales in North America spiked upwards while Japan and Europe remained level. The overall sales having had a higher average in every location before 1995 may be because video games could have been seen as something new before that year. Before 1995s could have also been when the first edition of what would later be franchises (such as Mario and Pokémon) were being released.

Plot 5.2) Average sales by genre



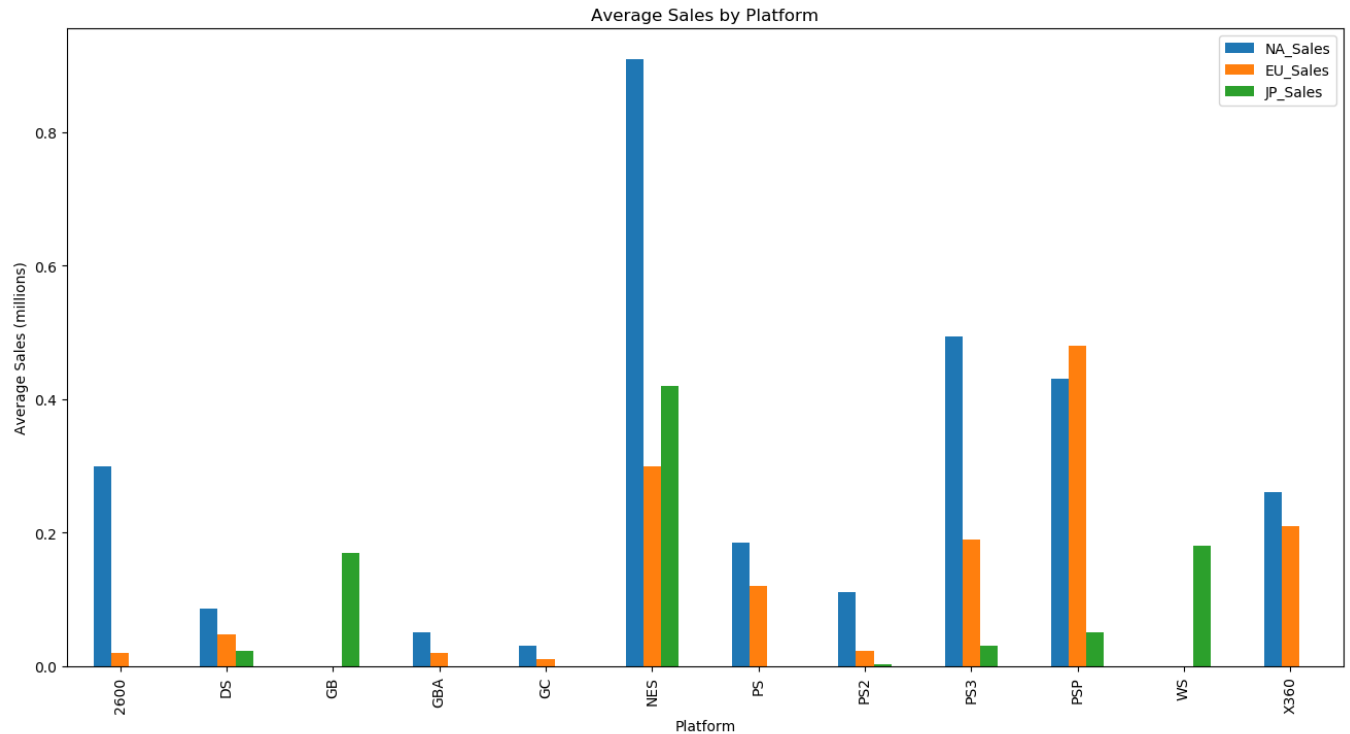
North America is again the biggest consumer across every genre. Both Platform and Shooter games are incredibly popular in North America, followed by Sports, Racing, Fighting, and Action. Both Platform and Shooter games are the oldest types of games which may be a factor for why both genres are so popular. Examples of Platform games include Donkey Kong, Mario Bros, Sonic, Pac-man and more. Europe matches the chart for North America though sales are not as high. In contrast to North America, Shooter games are the least popular genre in Japan. My hypothesis is that guns are a big part of citizen rights in the United States, while guns are very restricted in Japan. In contrast to both North America and Europe, Japan is very low on sales in racing but higher than the other two locations with role-playing games.

Plot 5.3) Average sales by publisher



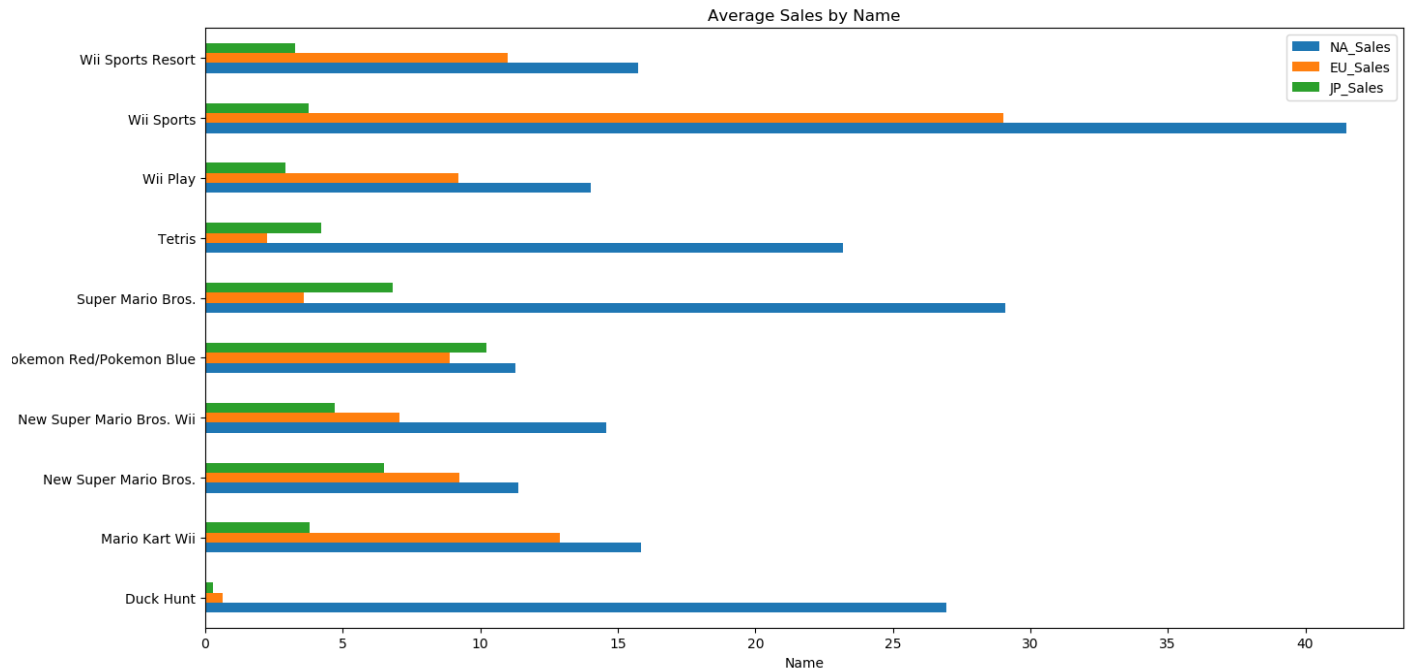
The popularity of publishers per country varies greatly. North American consumes the most in the majority of the publishers such as Parker Bros, Konami Digital Entertainment, Aspyr, and many others. Europe dominates in only two categories; Electronic Arts and Sega, despite Sega being a Japanese developer and publisher. In this specific bar graph of randomly picked publishers, Japan only uses three of the options and has the highest average sales in Nintendo. Europe has no publishers that are unique to that location. Meanwhile, the publishers Aspyr and Take-Two Initiative are unique to North American while Namco Bandai Games is unique to Japan.

Plot 5.4) Average sales by platform



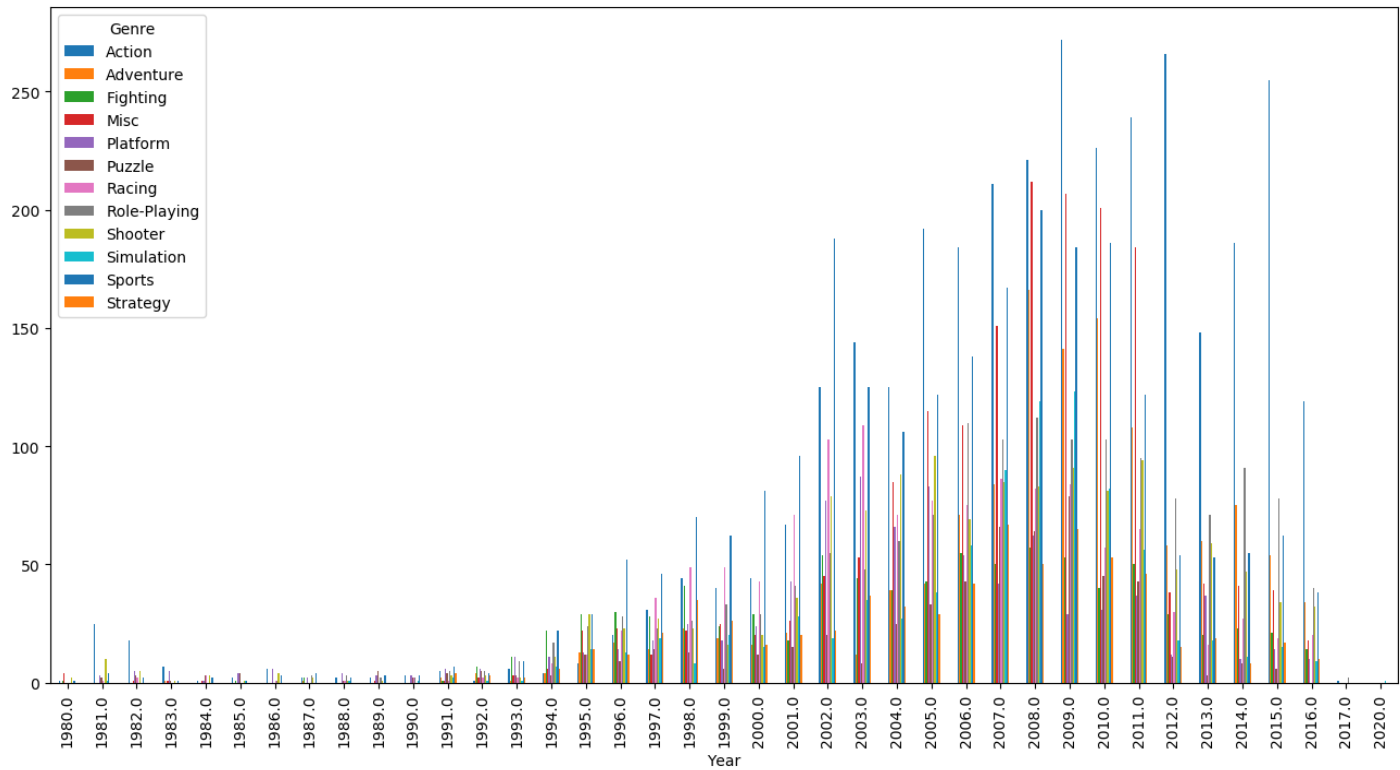
In nearly every platform listed, with the exception of GB, PSP, and WS, North America is the biggest consumer. If Europe contains a specific platform, so does North America. The only platform that Europe consumes more than North America is the PSP. In general, Europe consumes these specific platforms more than Japan, with the exception of NES. There are a few platforms, such as GB and WS, which solely garner money in Japan. The older video game platforms, such as the NES (Nintendo) and the PSP (PlayStation) have collected the most money. Physically smaller devices such as the GBA (Gameboy) and DS (Nintendo DS) in general have earned less than physically larger devices such as the X360 and PlayStation.

Plot 5.5) Average sales by name



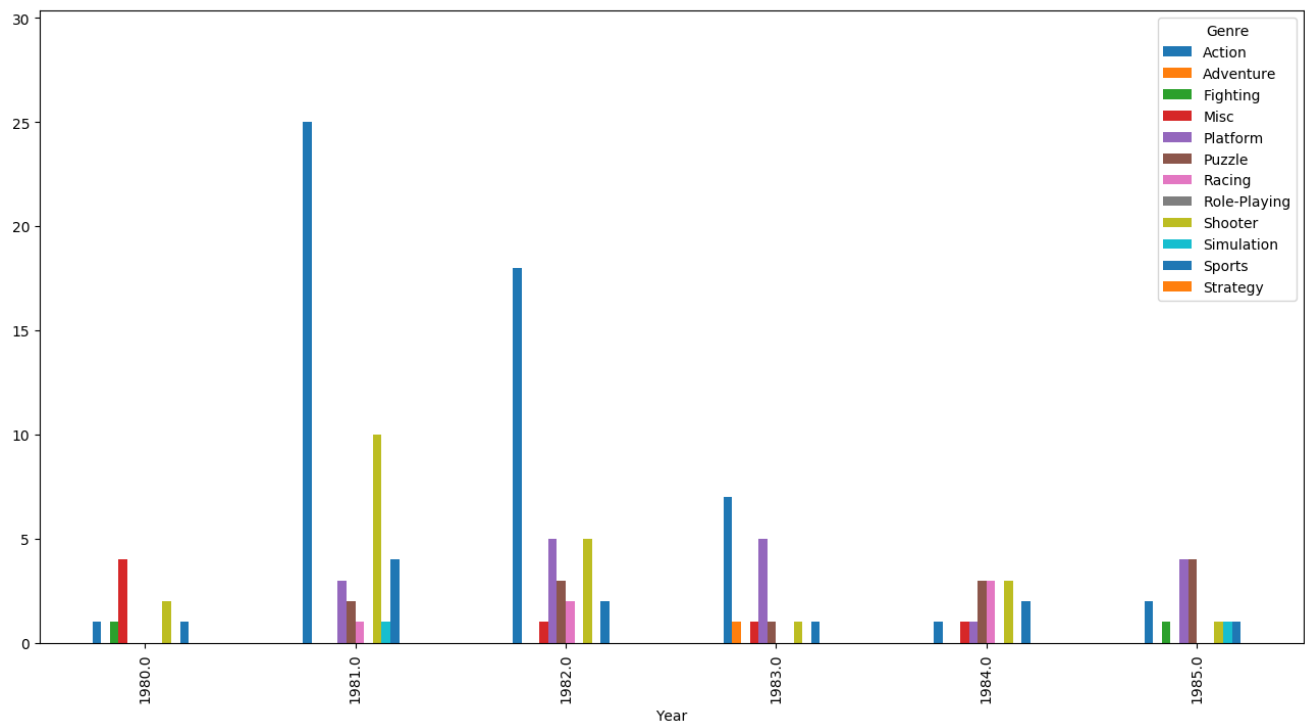
The top ten games according to this research study include five Wii games, four Mario games, two older games (Duck Hunt and Tetris both released in 1984), as well as the Pokémon Red/ Pokémon Blue. Pokémon Red/ Blue is the first installment of the Pokémon series. Every game is most popular in terms of monetary value in North America, despite some of the games such as Tetris and Pokémon Red/ Blue first being released in Japan. Europe consumes every top 10 video games listed more than Japan, with the exception of Tetris, Super Mario Bros, and Pokémon Red/ Blue. There are a few games such as Duck Hunt, Super Mario Bros, and Tetris that were far less popular in either Japan or Europe than North America.

Plot 6) Genre by Year



Based on an overall view of the genre by year graph, action and sports video games have been the most popular genres over every year from 1980 to 2016.

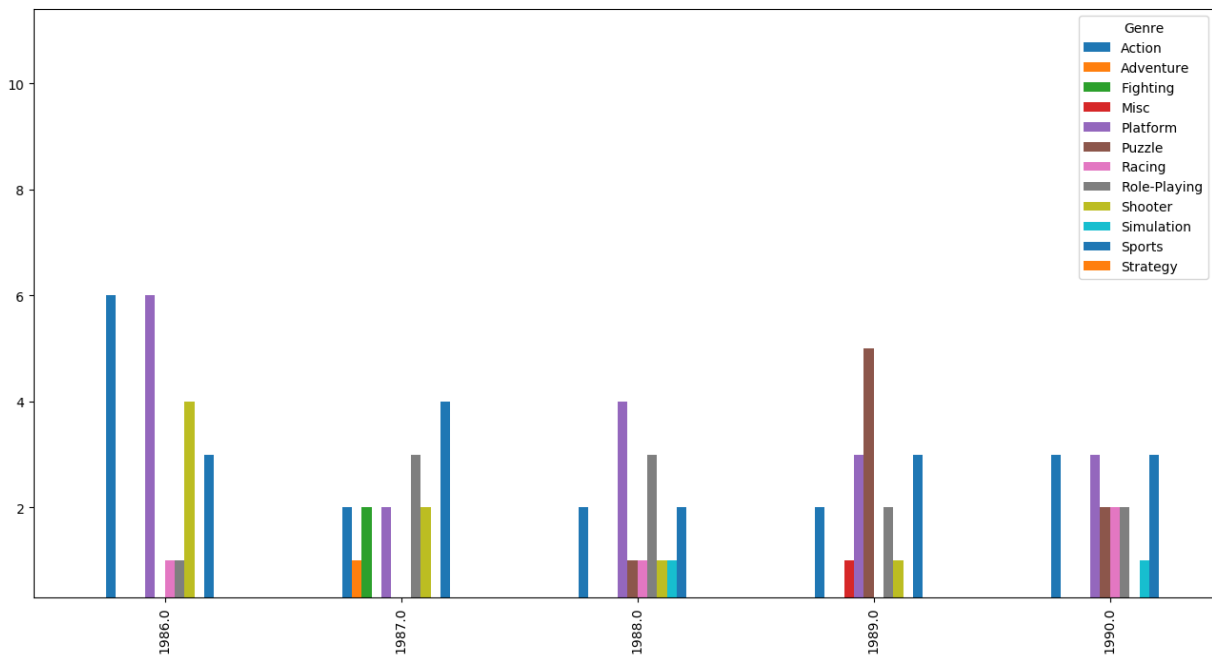
Plot 6.1) Genre by Year (1980-1985)



Note: The word “popular” in this context is used to refer to how much a video game genre is released per year.

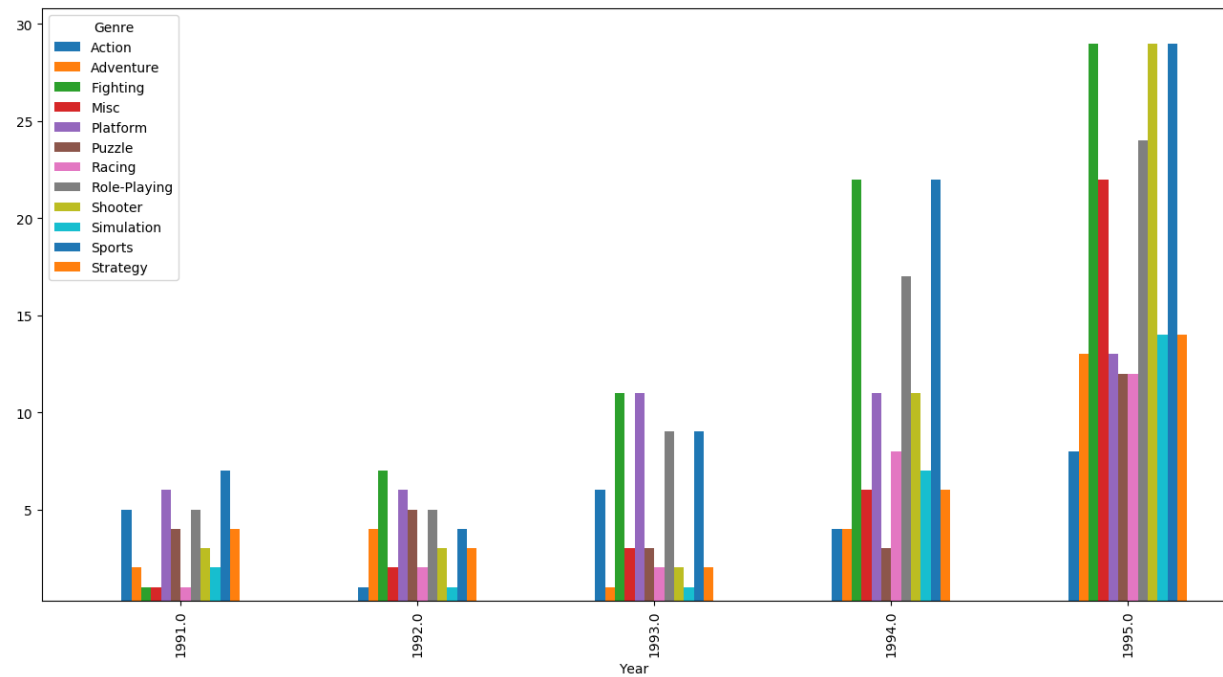
The first few video game genres released included shooter, action, sports, fighting, and miscellaneous. Action and shooter games were the most popular until 1984 when racing, puzzle, and platform grew in popularity.

Plot 6.2) Genre by Year (1986-1990)



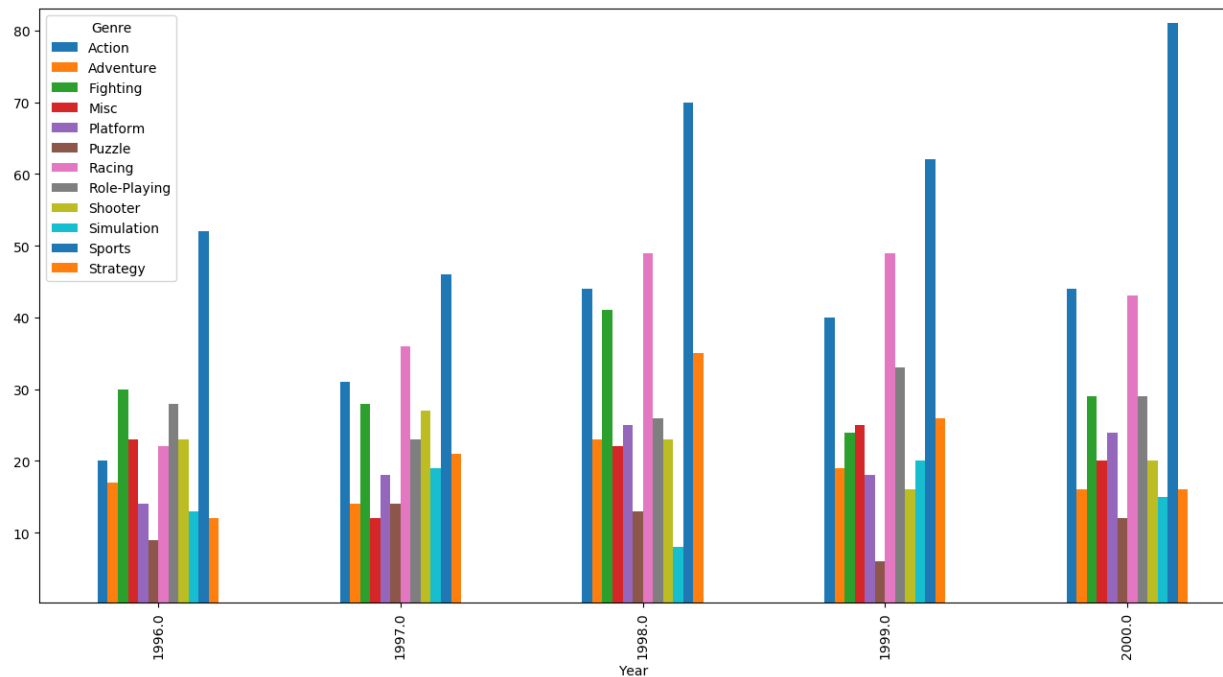
Action and platform games competed for the most popular position throughout 1986-1990 though in 1989, puzzle games became the most popular.

Plot 6.3) Genre by Year (1991-1995)



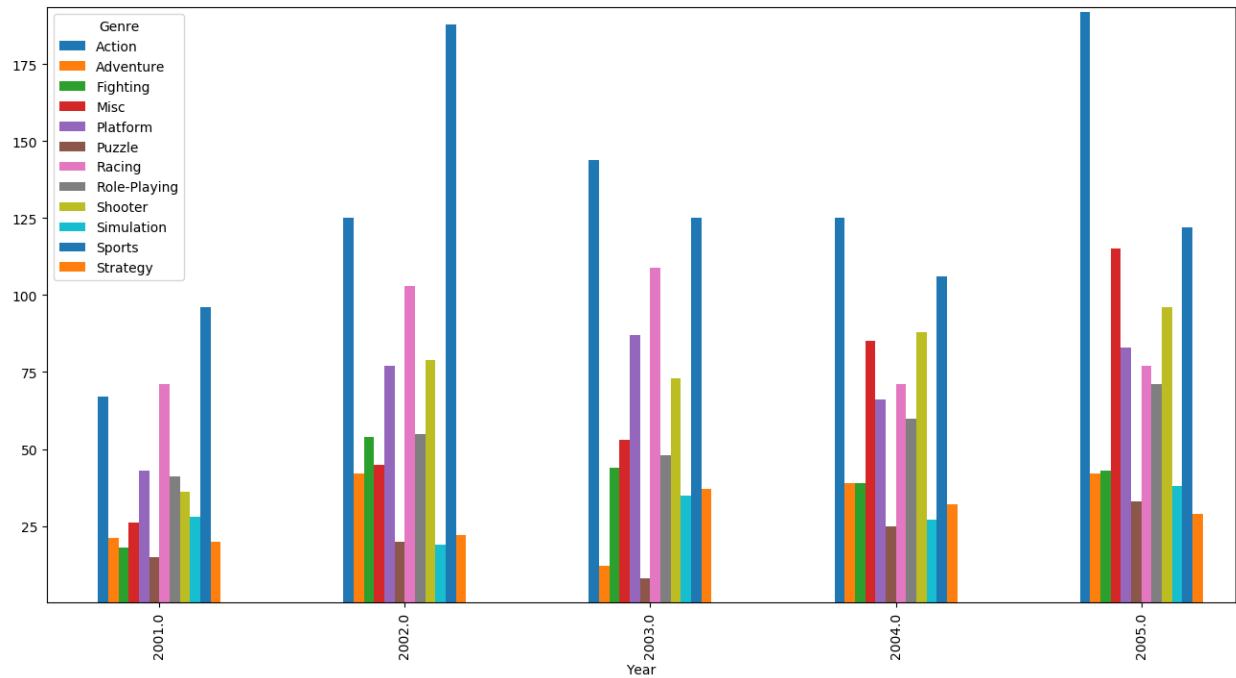
1991 was when all of the genres started being produced in the same year. Starting in 1992, fighting games grew quickly in popularity. By 1995, sports, shooter, and fighting games became the most popular. Role-playing games also grew consistently more popular throughout 1991-1995.

Plot 6.4) Genre by Year (1995-2000)



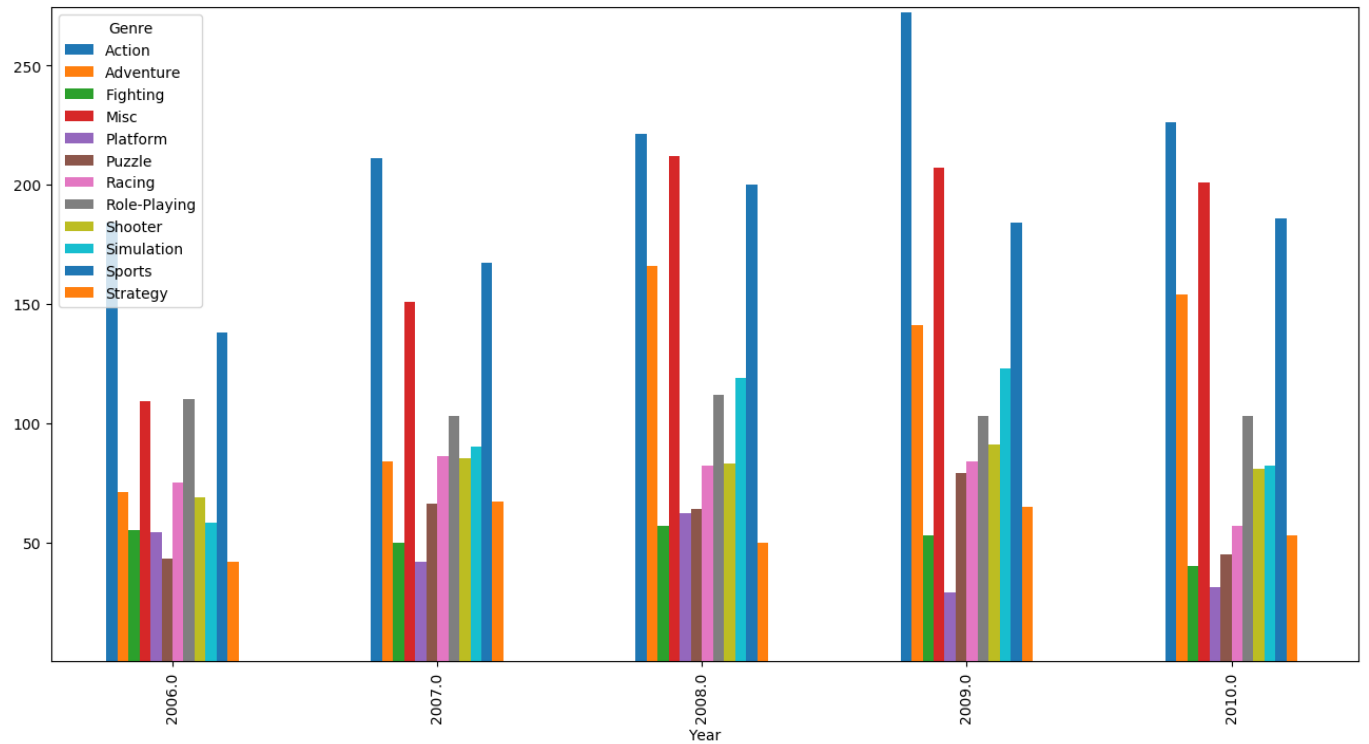
From 1996-2000, sports games were consistently the most popular games. Racing games also grew in popularity over the course of these years.

Plot 6.5) Genre by Year (2001-2005)



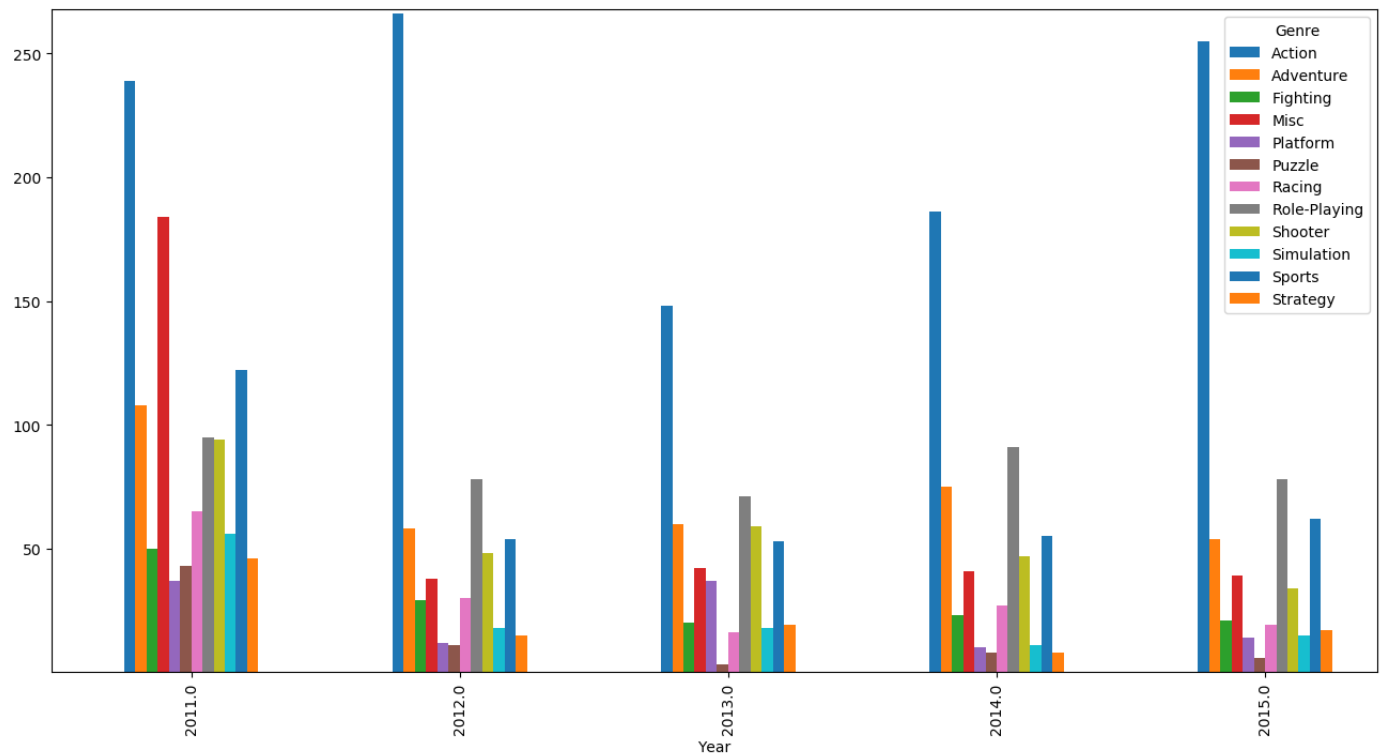
Action games and sports games remained consistently most popular throughout 2001-2005, with action garnering more popularity by 2005. The popularity of both miscellaneous and shooter games steadily grew over the course of these years as well. Puzzle and simulation games remained the least popular games.

Plot 6.6) Genre by Year (2006-2010)



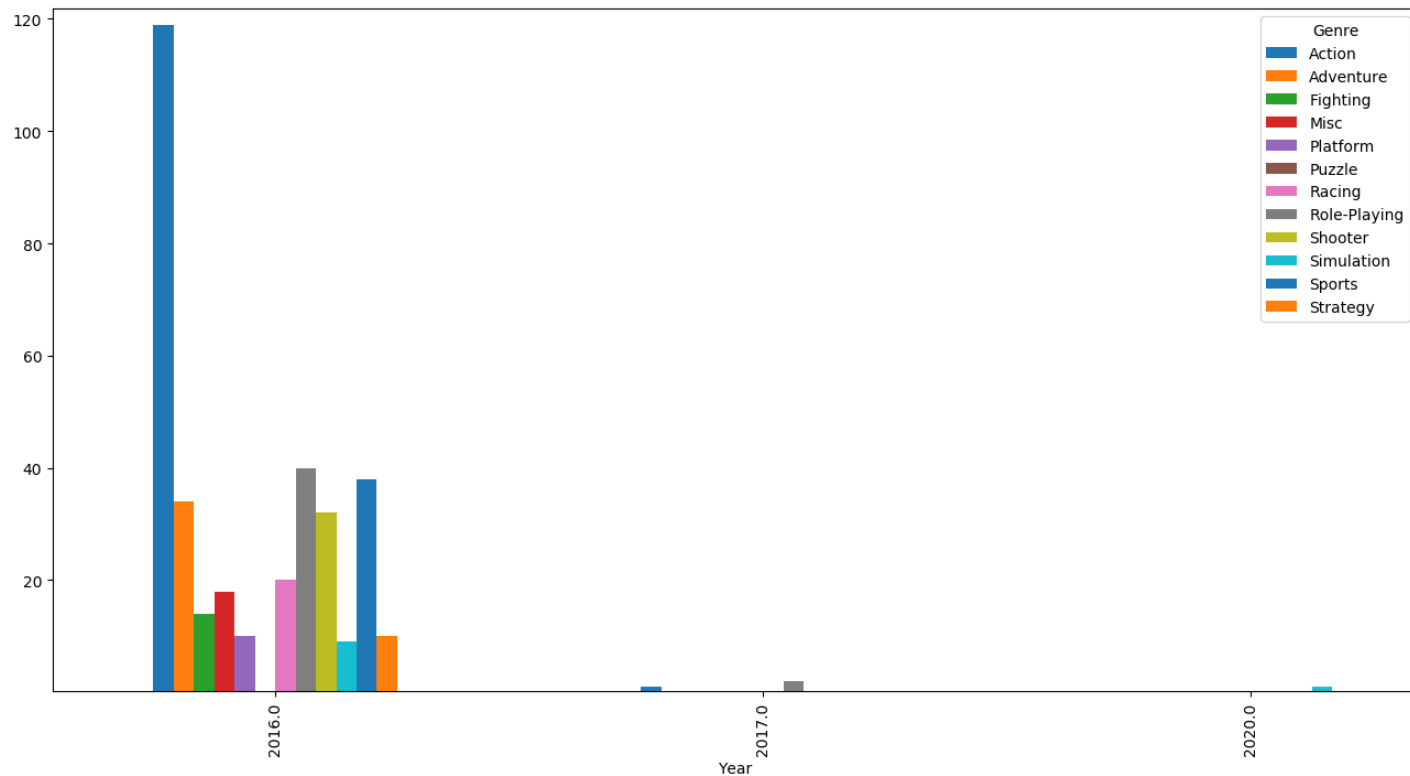
Action, sports, and miscellaneous games are the most popular throughout 2006-2010.

Plot 6.2) Genre by Year (2011-2015)



The popularity of action games grew rapidly throughout 2011-2015. Meanwhile, the popularity of platform and puzzle games decreased significantly.

Plot 6.2) Genre by Year (2016-2020)



Action games have remained the most popular in 2016. Role-playing games grew in popularity past sports and adventure games.

In comparison to other graphs:

Despite only having been the most popular genre in 1985, the platform genre earned the most in North American sales according to the Average Sales by genre graph. Role-playing games have never been the most popular by year but earn the most in Japan compared to Europe and North America. Despite the action genre becoming the most popular genre by 2016, its average sales are less than fighting, platform, racing, shooter, and sports. Despite shooter games being less popular than sports games for the majority of the years listed, shooter games earn more than sports games in average sales in North America and Europe. Despite adventure games being consistently more popular than puzzle games, puzzle games earn more than adventure games in Europe, Japan, and North America.

Acknowledgments to my group partner Nawal