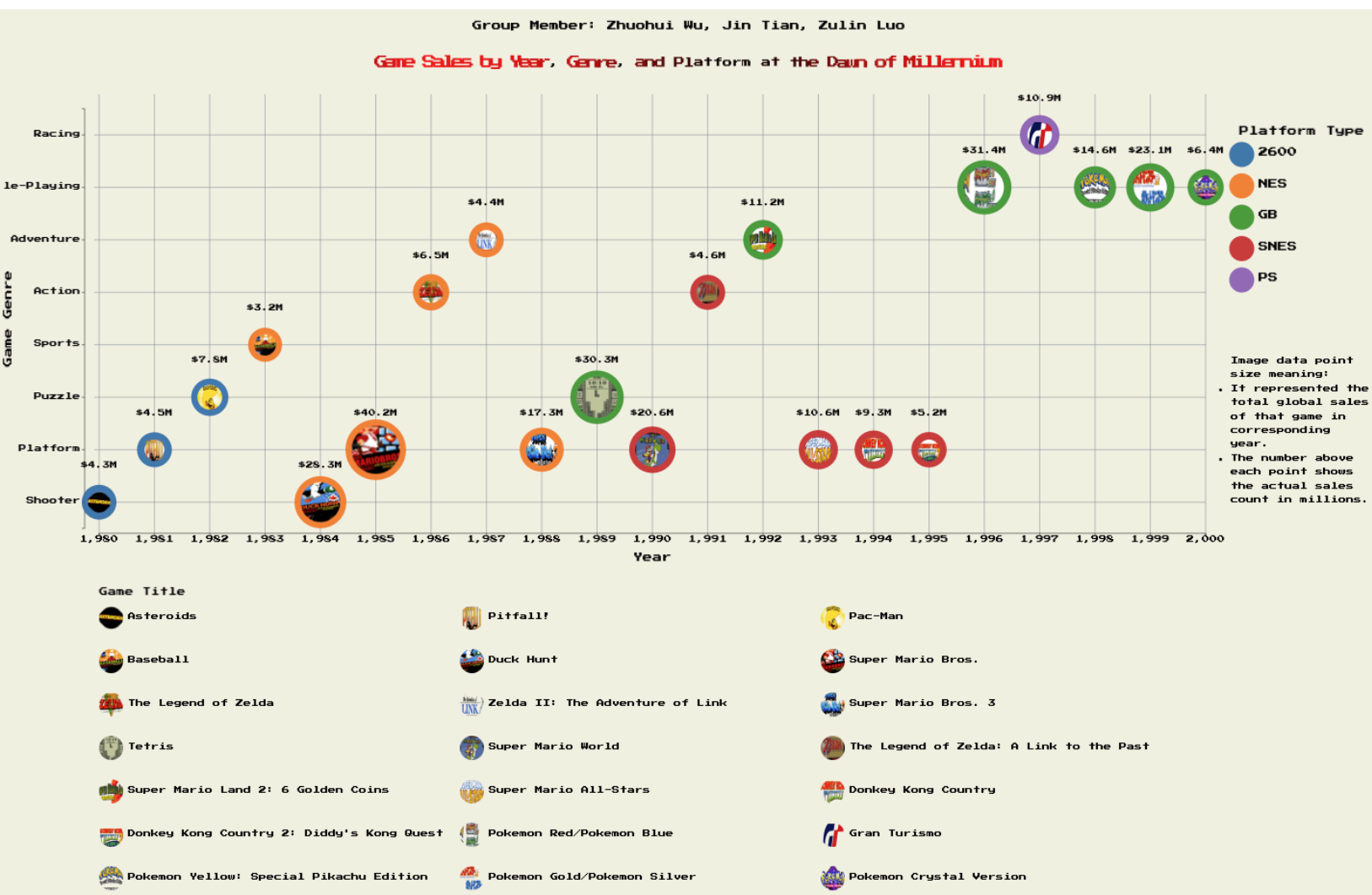


Group member: Zhuohui Wu, Jin Tian, Zulin Luo

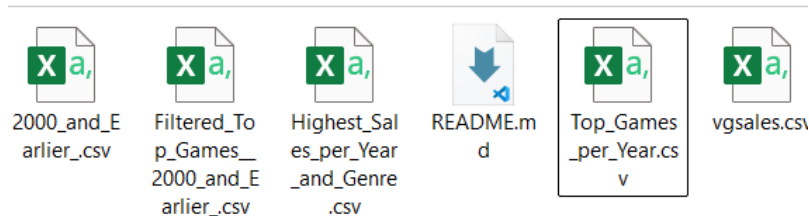
Screenshot of our plot



Group member: Zhuohui Wu, Jin Tian, Zulin Luo

Data Description

Our group decided to use the video game dataset to make out the visualization([Video Game Sales \(kaggle.com\)](https://www.kaggle.com/datasets/snoopyc/vgsales)). We decided to use game name, year, platform, genre, and sales to explain the difference between different video games in different years. The Dataset is extremely large



at first time. We decided to focus on video games before 2000. So, we use the tabular method to filter the data, only extracting the data we want. The first raw data we got is the vgsales.csv document. We filtered data in different ways including filtering the year before 2000 and sorting the game which has the highest sale in each year. And that is the data point we want to show in our final visualization. We want to let viewers know which games sell best each year before the Millennium.

Specific steps to clean and filter the data:

Year	Global_Sales
1,991	4.61
1,992	11.18
1,993	10.55
1,994	9.3
1,995	5.15
1,996	31.37
1,997	10.95
1,998	14.64
1,999	23.1
2,000	6.39

Step1: pull the file into tabular, add filter

Step 2: pull the year to the left and filter 1980 to 2000.

Step 3: aggregate year and global sales, set global sales to the max, which you will get each year's highest global sales

Step 4: After that, we can see all games are ranked by global sales.

Name	Platform	Year	Genre	Global_Sales
Super Mario Bros.	NES	1,985	Platform	40.24
Pokemon Red/Pokemon Blue	GB	1,996	Role-Playing	31.37
Tetris	GB	1,989	Puzzle	30.26
Duck Hunt	NES	1,984	Shooter	28.31
Pokemon Gold/Pokemon Silver	GB	1,999	Role-Playing	23.1
Super Mario World	SNES	1,990	Platform	20.61
Super Mario Land	GB	1,989	Platform	18.14
Super Mario Bros. 3	NES	1,988	Platform	17.28
Pokémon Yellow: Special Pikachu Edition	GB	1,998	Role-Playing	14.64
Super Mario 64	N64	1,996	Platform	11.89

Group member: Zhuohui Wu, Jin Tian, Zulin Luo

Step 5: Then we just need to extract each year's top global sales game.

Step 6: Finally, export data in a CSV file.

A	B	C	D	E	F	G	H	I	J	K	L	M	N
	Unnamed: Rank		Name	Platform	Year	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Other_Sale	Global_Sales	
0	258	259	Asteroids	2600	1980	Shooter	Atari	4	0.26	0	0.05	4.31	
1	239	240	Pitfall!	2600	1981	Platform	Activision	4.21	0.24	0	0.05	4.5	
2	89	90	Pac-Man	2600	1982	Puzzle	Atari	7.28	0.45	0	0.08	7.81	
3	421	422	Baseball	NES	1983	Sports	Nintendo	0.73	0.1	2.35	0.02	3.2	
4	9	10	Duck Hunt	NES	1984	Shooter	Nintendo	26.93	0.63	0.28	0.47	28.31	
5	1	2	Super Mari	NES	1985	Platform	Nintendo	29.08	3.58	6.81	0.77	40.24	
6	127	128	The Legend	NES	1986	Action	Nintendo	3.74	0.93	1.69	0.14	6.51	
7	251	252	Zelda II: Th	NES	1987	Adventure	Nintendo	2.19	0.5	1.61	0.08	4.38	
8	22	23	Super Mari	NES	1988	Platform	Nintendo	9.54	3.44	3.84	0.46	17.28	
9	5	6	Tetris	GB	1989	Puzzle	Nintendo	23.2	2.26	4.22	0.58	30.26	
10	18	19	Super Mari	SNES	1990	Platform	Nintendo	12.78	3.75	3.54	0.55	20.61	
11	231	232	The Legend	SNES	1991	Action	Nintendo	2.42	0.91	1.15	0.13	4.61	
12	50	51	Super Mari	GB	1992	Adventure	Nintendo	6.16	2.04	2.69	0.29	11.18	
13	57	58	Super Mari	SNES	1993	Platform	Nintendo	5.99	2.15	2.12	0.29	10.55	
14	71	72	Donkey Ko	SNES	1994	Platform	Nintendo	4.36	1.71	3	0.23	9.3	
15	187	188	Donkey Ko	SNES	1995	Platform	Nintendo	2.1	0.74	2.2	0.11	5.15	
16	4	5	Pokemon F	GB	1996	Role-Playi	Nintendo	11.27	8.89	10.22	1	31.37	
17	52	53	Gran Turis	PS	1997	Racing	Sony Comp	4.02	3.87	2.54	0.52	10.95	
18	30	31	Pokemon Y	GB	1998	Role-Playi	Nintendo	5.89	5.04	3.12	0.59	14.64	
19	12	13	Pokemon C	GB	1999	Role-Playi	Nintendo	9	6.18	7.2	0.71	23.1	
20	132	133	Pokemon C	GB	2000	Role-Playi	Nintendo	2.55	1.56	1.29	0.99	6.39	

Design Rationale

1. Scatterplot

- We created a scatterplot to represent the games with the highest global sales from 1980 to 2000 with their respective genres and platforms. This chart type can better help show both quantitative data and categorical data. We included a title, *Game Sales by Year, Genre, and Platform at the Dawn of Millennium*, to describe the purpose of the visualization. Also, we added labels to both the x and y axes to help the audience better understand the graph. To make the sizes of each data point more self-explanatory, we added the exact global sales amount (e.g. \$40.2M) labels above each point to facilitate the comparison. This graph aims to create an informative and interesting visualization of the highest global sales game by year, genre, and platform and show the trends over time.

2. Position

- The x-axis represents the year of the game's release. We used a linear scale to show the standard data points for each year. Since we only show the highest sales game, which is 1 data point, the linear scale can effectively show our data.
- The y-axis represents game genres. We used a band scale to more effectively show the nominal data of the genres. This scale can organize genres distinctly and enables easier comparison within the same genre.

3. Gridlines

- We added gridlines to both the x and y axes to enhance the readability and allow the audience to trace data points more easily.

4. Marks

- We used circular images of the game logos as our marks (e.g., The Legend of Zelda, Super Mario Bros). We downloaded the PNG images online, and each data point corresponds to the highest global sales game for each year. We want to make the data points more visually appealing and provide direct information in terms of what game this point represents.

5. Size of Marks

- The size of each mark is determined by the total global sales count. We used a linear scale to select a size range between 40 and 100 pixels. This scale allows for a clearer contrast in sizes, in which the lowest global sales correspond to a 40-pixel size, and the highest sales are represented by a 100-pixel size. This will make it easier for the audience to differentiate in sales count at a glance.

6. Stroke Color

- The stroke color of the image circles (marks) is determined by an ordinal scale using "d3.schemeCategory10". This color scheme is used to differentiate between the different game platforms. We chose a stroke width of 5 pixels to ensure that each circle stroke stands out clearly against the background.

7. Legends

- We included two legends: one legend explains which colors correspond to each released platform, while another legend explains which logo image corresponds to each game title. This is to help our audience better understand the color representation and the game name in case they are unfamiliar with video game icons.

8. Styles

- We chose "Press Start 2P" as our font as it mimics the font of the video game style in the Millennium years. We added this to make our aesthetic more interesting and bonded to our topic of the graph and dataset.
- We also added some minor animation experiments like the neon title and the flashing strokes to make the whole visualization more appealing and unique. But even without these minor animations, our graph is still very readable and complete.

9. Trade-offs

- The decision to use images rather than simple shapes enhances the appealingness of our graph, However, this might lead to a less easy comparison of the sizes and potential overplotting in crowded areas. To solve this, we added strokes and the actual sales count label to help the audience comprehend. Also, we only included one data point for each year to avoid including too much data. We focused only on the games with the highest global sales.
- Since we only included the games with the highest global sales, some less known or sold games might not be able to be shown on the graph.

Visualization Story

The scatterplot shows the overall trends of the most-sold games in each year between 1980 to 2000 and their corresponding information.

Popularity of Game Genres:

Between 1980 and 1995, shooter, platform, and puzzle games maintained consistent popularity. Among these genres, platform games appeared to be the most favored ones for game players in these years. Action and adventure games also showed significant sales in various years. For instance, games like Super Mario Bros. and The Legend of Zelda had significant sales. After 1996, there was a notable shift as role-playing games took over the market. This genre became the leading genre in the latter part of the decade. This change reflects the evolving player preferences and the introduction of more immersive gameplay experiences.

Platform Evolution:

The visualization also shows some trends in console preferences over the years. The 2600 video game consoles dominated the market before 1982. Between 1983 and 1988, the NES (Nintendo Entertainment System) console began to lead. This console-type revolutionized game players' gaming experience and started to increase in its market share. Afterward, the SNES (Super Nintendo Entertainment System) console gained popularity between 1990 and 1995. As technology advanced, other platforms like the GB (Game Boy) and PS (PlayStation) emerged, and they became popular choices after 1996. This shift indicates a diversification in gamer preferences about console types over the years.

Sales Figures:

Notably, Super Mario Bros. stands out with an outstanding \$40.2 million in global sales, which makes it the highest-global-selling game in the dataset. We were surprised by this sales count and how it might lead to the future trend in platform games. This figure is particularly impressive given the time in which it was released, and this definitely emphasizes its significant cultural impact and widespread appeal.

Team contributions

Zhuohui Wu: Collaborated closely with the team to guide the selection of the data's x-axis and y-axis, laying the groundwork for the scatter plot's foundational code. Enhanced the visualization by incorporating dynamic effects and pixel-style fonts, adding a unique touch to the chart. Added a game-themed legend to improve the plot's interpretability and visual appeal. Additionally, provided support to team members by troubleshooting coding challenges and offering solutions, ensuring the successful implementation of the project.

Group member: Zhuohui Wu, Jin Tian, Zulin Luo

Tian Jin: Actively discussing data use with teammates. Creating axis labels, and arranging the x and y-axis labels to make them look more clear in shape and font. Adding platform legend with color to differentiate each game with different platforms. Helping to find bugs, discussing solutions with teammates, and possible alternatives. Collecting, sorting, and cropping all video game icons to make all icons in the same scale and circle shape. Inserting all icons into code. Also, help team members generate data descriptions and visualization stories.

Zulin Luo: Worked with the team to generate the graph and adjust the design to make it readable and clear. Provided many design ideas and experimented with different styles. Proactively discussed ideas with the team and led meetings to work on the coding together. Drafted the image mark codes and tried with other data columns and images. Added the data labels to the graph. Adjusted the legend and added a description of the data size to the side of the graph. Wrote the design rational report with details. Fixed errors and revised the visualization story to make it more clear and comprehensive.