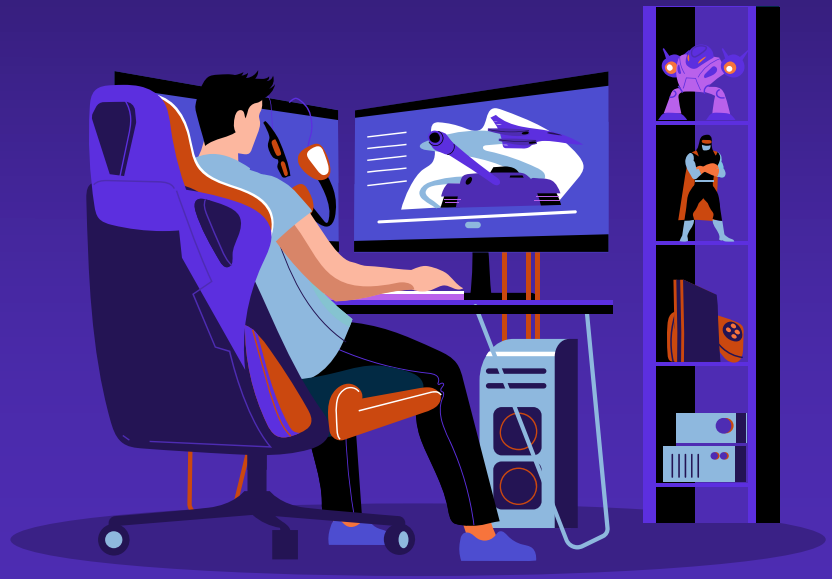
Video Game Sales

Data Analysis

By Nigel



Summary

Over the past couple of decades, video games has been rising in popularity among youths, and even adults throughout the world.
In this project, a data analysis will be done on the sales of video games to evaluate the sales of video games throughout different countries.

Table of contents

01 Problem Statement

02 Data Preparation

03 Data Analysis

04 Insights & Recommendation

05 Project stages

06 Our team

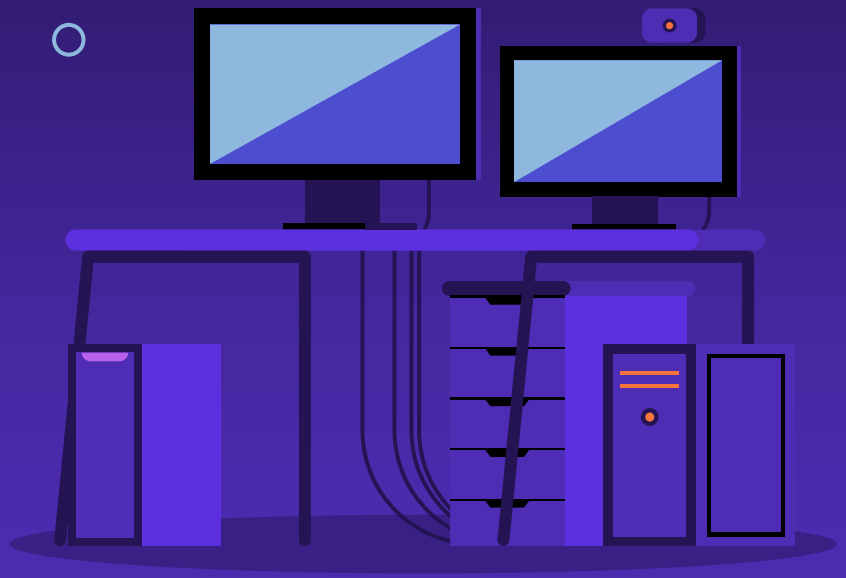
01

Problem Statement

You can enter a subtitle here if you need it

Problem Statement

- Which genres are more popular? Do they differ in each region?
- Which video game platforms are more popular?
- Which specific games are most popular?
- How can video gaming companies boost their sales using the data analyzed?



02

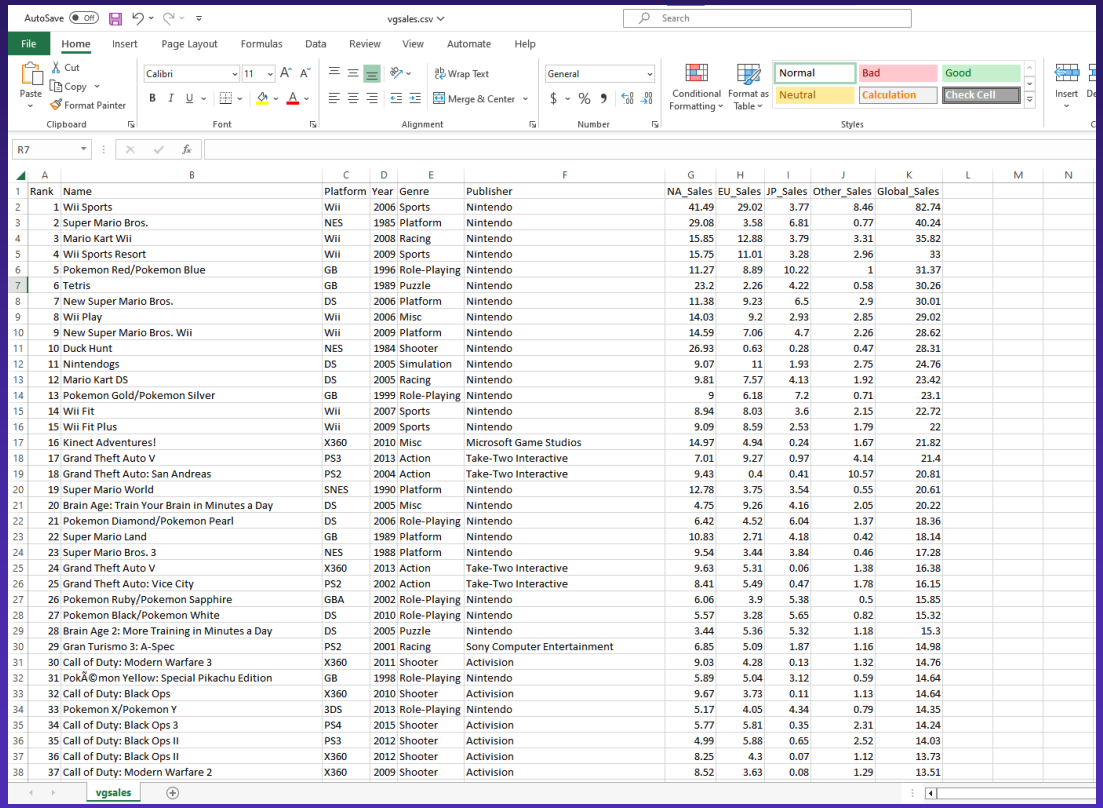
Data Preparation & Processing

You can enter a subtitle here if you need it

Data Preparation

- The dataset used is from Kaggle - https://www.kaggle.com/datasets/gregorut/videogamesales
- As mentioned by the author, the dataset was scrapped from the website www.vgchartz.com, using the Python script BeautifulSoup.
- The fields of the dataset include:
  - Rank - Ranking of overall sales
  - Name - The games name
  - Platform - Platform of the games release (i.e. PC,PS4, etc.)
  - Year - Year of the game's release
  - Genre - Genre of the game
  - Publisher - Publisher of the game
  - NA_Sales - Sales in North America (in millions)
  - EU_Sales - Sales in Europe (in millions)
  - JP_Sales - Sales in Japan (in millions)
  - Other_Sales - Sales in the rest of the world (in millions)
  - Global_Sales - Total worldwide sales

Data Processing (Excel Overview)



AutoSave Off vgsales.csv Search

File Home Insert Page Layout Formulas Data Review View Automate Help

Paste Cut Copy Format Painter Clipboard | Calibri 11 Font | Wrap Text Merge & Center Alignment | General Number | Conditional Formatting Format as Table | Normal Bad Good Neutral Calculation Check Cell Styles | Insert

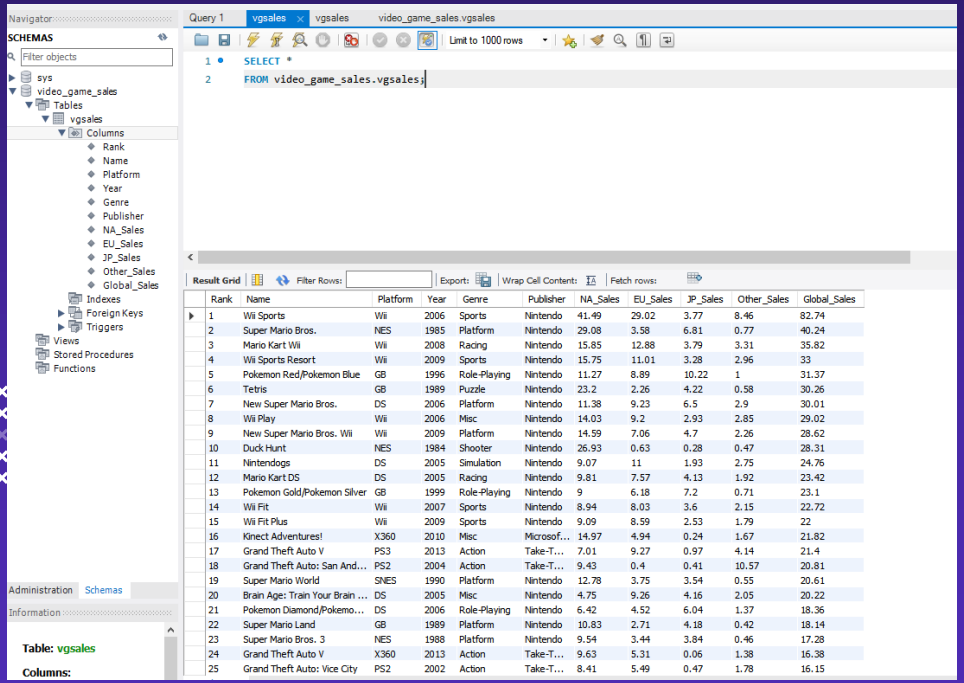
R7

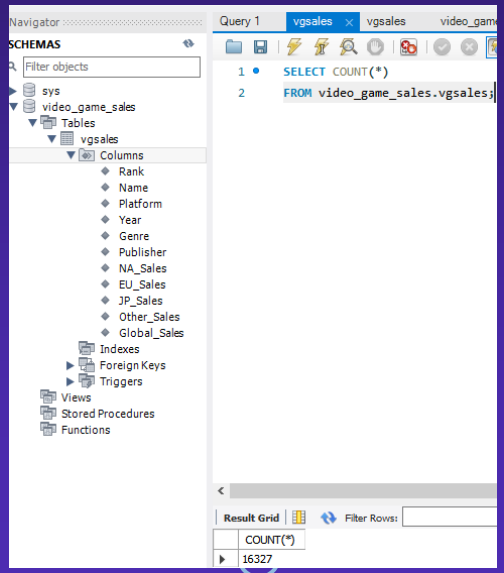
	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Rank	Name	Platform	Year	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales			
2	1	Wii Sports	Wii	2006	Sports	Nintendo	41.49	29.02	3.77	8.46	82.74			
3	2	Super Mario Bros.	NES	1985	Platform	Nintendo	29.08	3.58	6.81	0.77	40.24			
4	3	Mario Kart Wii	Wii	2008	Racing	Nintendo	15.85	12.88	3.79	3.31	35.82			
5	4	Wii Sports Resort	Wii	2009	Sports	Nintendo	15.75	11.01	3.28	2.96	33			
6	5	Pokemon Red/Pokemon Blue	GB	1996	Role-Playing	Nintendo	11.27	8.89	10.22	1	31.37			
7	6	Tetris	GB	1989	Puzzle	Nintendo	23.2	2.26	4.22	0.58	30.26			
8	7	New Super Mario Bros.	DS	2006	Platform	Nintendo	11.38	9.23	6.5	2.9	30.01			
9	8	Wii Play	Wii	2006	Misc	Nintendo	14.03	9.2	2.93	2.85	29.02			
10	9	New Super Mario Bros. Wii	Wii	2009	Platform	Nintendo	14.59	7.06	4.7	2.26	28.62			
11	10	Duck Hunt	NES	1984	Shooter	Nintendo	26.93	0.63	0.28	0.47	28.31			
12	11	Nintendogs	DS	2005	Simulation	Nintendo	9.07	11	1.93	2.75	24.76			
13	12	Mario Kart DS	DS	2005	Racing	Nintendo	9.81	7.57	4.13	1.92	23.42			
14	13	Pokemon Gold/Pokemon Silver	GB	1999	Role-Playing	Nintendo	9	6.18	7.2	0.71	23.1			
15	14	Wii Fit	Wii	2007	Sports	Nintendo	8.94	8.03	3.6	2.15	22.72			
16	15	Wii Fit Plus	Wii	2009	Sports	Nintendo	9.09	8.59	2.53	1.79	22			
17	16	Kinect Adventures!	X360	2010	Misc	Microsoft Game Studios	14.97	4.94	0.24	1.67	21.82			
18	17	Grand Theft Auto V	PS3	2013	Action	Take-Two Interactive	7.01	9.27	0.97	4.14	21.4			
19	18	Grand Theft Auto: San Andreas	PS2	2004	Action	Take-Two Interactive	9.43	0.4	0.41	10.57	20.81			
20	19	Super Mario World	SNES	1990	Platform	Nintendo	12.78	3.75	3.54	0.55	20.61			
21	20	Brain Age: Train Your Brain in Minutes a Day	DS	2005	Misc	Nintendo	4.75	9.26	4.16	2.05	20.22			
22	21	Pokemon Diamond/Pokemon Pearl	DS	2006	Role-Playing	Nintendo	6.42	4.52	6.04	1.37	18.36			
23	22	Super Mario Land	GB	1989	Platform	Nintendo	10.83	2.71	4.18	0.42	18.14			
24	23	Super Mario Bros. 3	NES	1988	Platform	Nintendo	9.54	3.44	3.84	0.46	17.28			
25	24	Grand Theft Auto V	X360	2013	Action	Take-Two Interactive	9.63	5.31	0.06	1.38	16.38			
26	25	Grand Theft Auto: Vice City	PS2	2002	Action	Take-Two Interactive	8.41	5.49	0.47	1.78	16.15			
27	26	Pokemon Ruby/Pokemon Sapphire	GBA	2002	Role-Playing	Nintendo	6.06	3.9	5.38	0.5	15.85			
28	27	Pokemon Black/Pokemon White	DS	2010	Role-Playing	Nintendo	5.57	3.28	5.65	0.82	15.32			
29	28	Brain Age 2: More Training in Minutes a Day	DS	2005	Puzzle	Nintendo	3.44	5.36	5.32	1.18	15.3			
30	29	Gran Turismo 3: A-Spec	PS2	2001	Racing	Sony Computer Entertainment	6.85	5.09	1.87	1.16	14.98			
31	30	Call of Duty: Modern Warfare 3	X360	2011	Shooter	Activision	9.03	4.28	0.13	1.32	14.76			
32	31	PokÃ©mon Yellow: Special Pikachu Edition	GB	1998	Role-Playing	Nintendo	5.89	5.04	3.12	0.59	14.64			
33	32	Call of Duty: Black Ops	X360	2010	Shooter	Activision	9.67	3.73	0.11	1.13	14.64			
34	33	Pokemon X/Pokemon Y	3DS	2013	Role-Playing	Nintendo	5.17	4.05	4.34	0.79	14.35			
35	34	Call of Duty: Black Ops 3	PS4	2015	Shooter	Activision	5.77	5.81	0.35	2.31	14.24			
36	35	Call of Duty: Black Ops II	PS3	2012	Shooter	Activision	4.99	5.88	0.65	2.52	14.03			
37	36	Call of Duty: Black Ops II	X360	2012	Shooter	Activision	8.25	4.3	0.07	1.12	13.73			
38	37	Call of Duty: Modern Warfare 2	X360	2009	Shooter	Activision	8.52	3.63	0.08	1.29	13.51			

vgsales

Data Processing (SQL Overview)

- As this is a large dataset (16,327 rows), we will mainly be using SQL to process the data (mySQL).





Data Processing (mySQL)

- REMOVING DUPLICATES
  - The following query helps to identify duplicates in the ‘Name’, ‘Platform’ and ‘Year’ column:

```
SELECT
Name, COUNT(Name),
Platform, COUNT(Platform)
Year, COUNT(Year)

FROM
video_game_sales.vgsales

GROUP BY
        Name,
        Platform,
        Year

HAVING
        (COUNT(Name) > 1) AND
        (COUNT(Platform) > 1)
```

From the results, we can see that ‘Madden NFL 13’ has a duplicate:

Result Grid | Filter Rows: | Export: | Wrap

Name	COUNT(Name)	Platform	Year	COUNT(Year)
Madden NFL 13	2	PS3	2	2

Data Processing (mySQL)

- REMOVING DUPLICATES

- Upon further inspection, we can see that row 16130 is the duplicate row;

```
SELECT *

FROM video_game_sales.vgsales

WHERE Name = 'Madden NFL 13'
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Rank	Name	Platform	Year	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales
507	Madden NFL 13	X360	2012	Sports	Electronic Arts	2.53	0.16	0	0.17	2.86
604	Madden NFL 13	PS3	2012	Sports	Electronic Arts	2.11	0.23	0	0.22	2.56
3732	Madden NFL 13	Wii	2012	Sports	Electronic Arts	0.51	0	0	0.03	0.54
5590	Madden NFL 13	PSV	2012	Sports	Electronic Arts	0.3	0	0	0.03	0.32
6794	Madden NFL 13	WiiU	2012	Sports	Electronic Arts	0.22	0	0	0.02	0.24
16130	Madden NFL 13	PS3	2012	Sports	Electronic Arts	0	0.01	0	0	0.01

- Deleting row 16130;

```
DELETE FROM video_game_sales.vgsales

WHERE Name = 'Madden NFL 13' AND Global_Sales = 0.01
```

- Double checking;

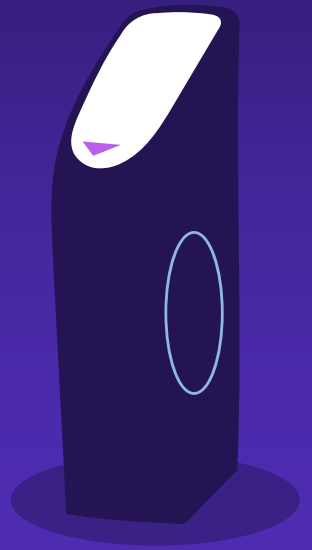
```
SELECT *
FROM video_game_sales.vgsales
WHERE Name = 'Madden NFL 13'
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Rank	Name	Platform	Year	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales
507	Madden NFL 13	X360	2012	Sports	Electronic Arts	2.53	0.16	0	0.17	2.86
604	Madden NFL 13	PS3	2012	Sports	Electronic Arts	2.11	0.23	0	0.22	2.56
3732	Madden NFL 13	Wii	2012	Sports	Electronic Arts	0.51	0	0	0.03	0.54
5590	Madden NFL 13	PSV	2012	Sports	Electronic Arts	0.3	0	0	0.03	0.32
6794	Madden NFL 13	WiiU	2012	Sports	Electronic Arts	0.22	0	0	0.02	0.24

03

Data Analysis



Data Analysis - Sales

- Looking at the **average sales** for each country:

```
SELECT
        AVG(NA_Sales) AS NA, AVG(EU_Sales) AS EU,
        AVG(JP_Sales) AS JP, AVG(Global_Sales) AS Global

FROM
        video_game_sales.vgsales
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

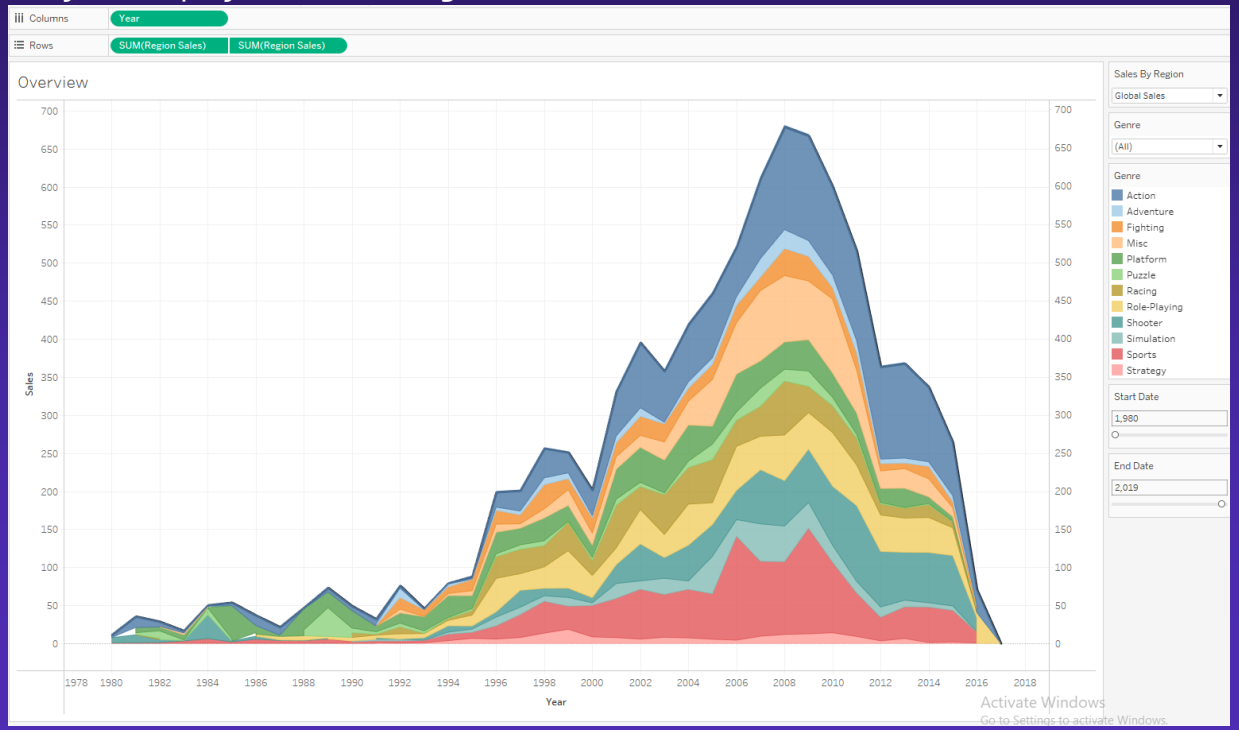
NA	EU	JP	Global
0.2654312140144755	0.14756278329048997	0.07866593164277245	0.5402639960799245

Note: Sales are in millions of copies

- North America has the highest average sales (260,000), followed by Europe (140,000) then Japan (78,000).

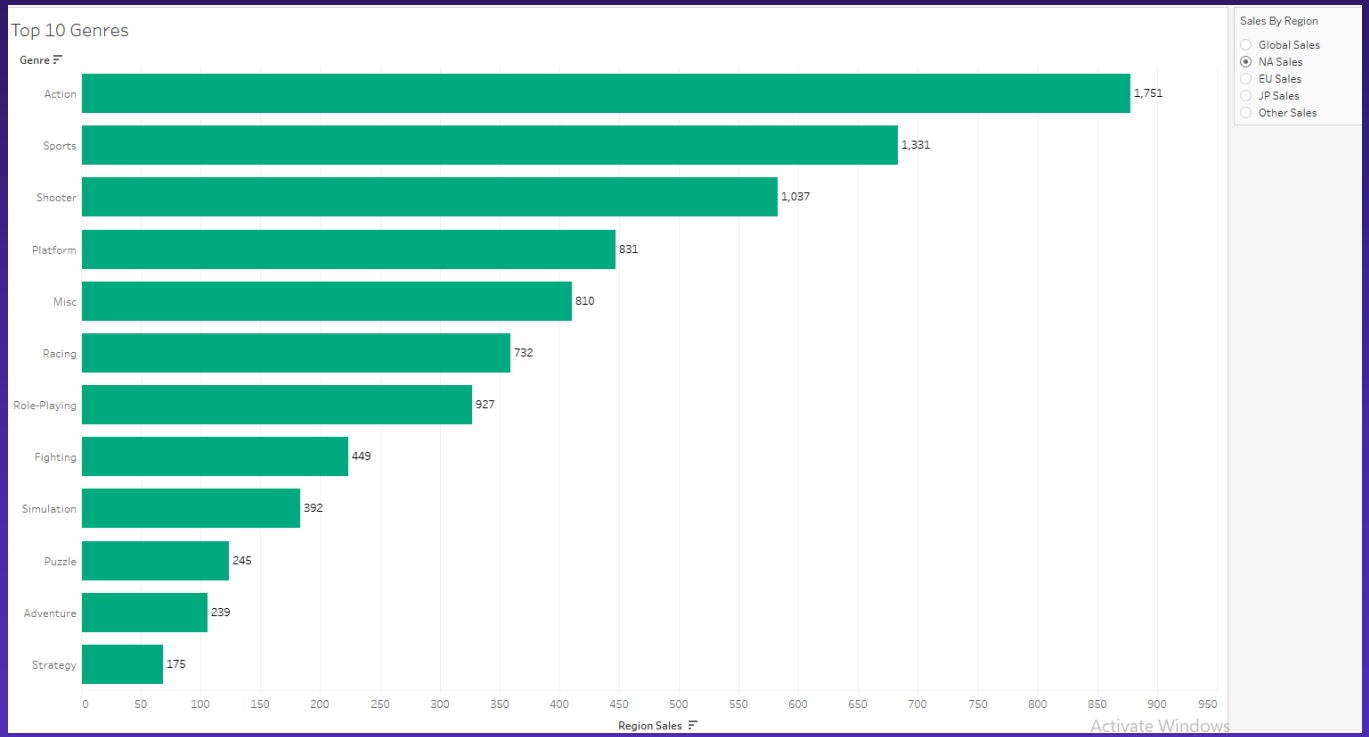
Data Analysis - Sales

- Using Tableau, we can visualize the sales of video games by its Region, Genre & Year.
- Tableau Live Link: https://public.tableau.com/views/VideoGameSales_16715481770100/Overview?:language=en-US&publish=yes&:display_count=n&:origin=viz_share_link



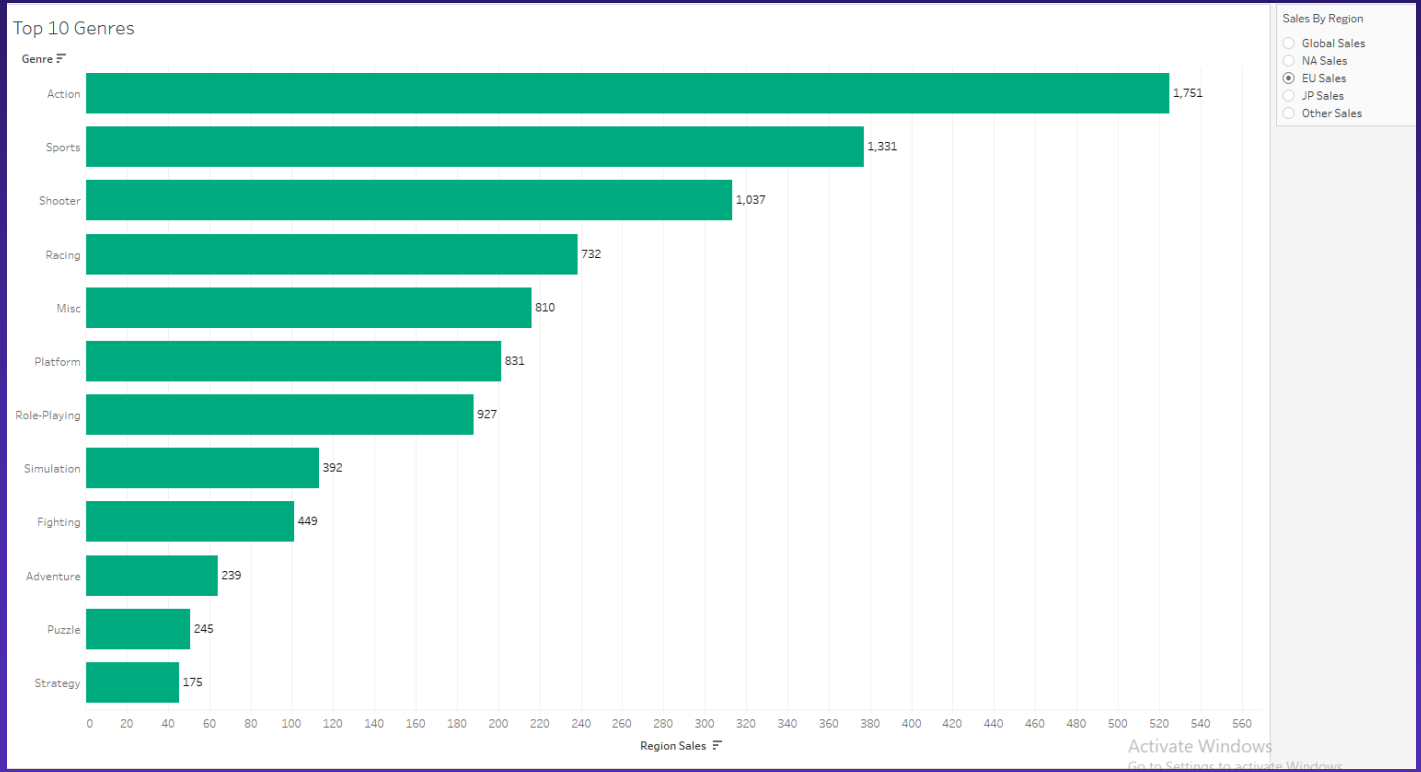
Data Analysis – Genre (NA)

- Top 10 Genres in North America:



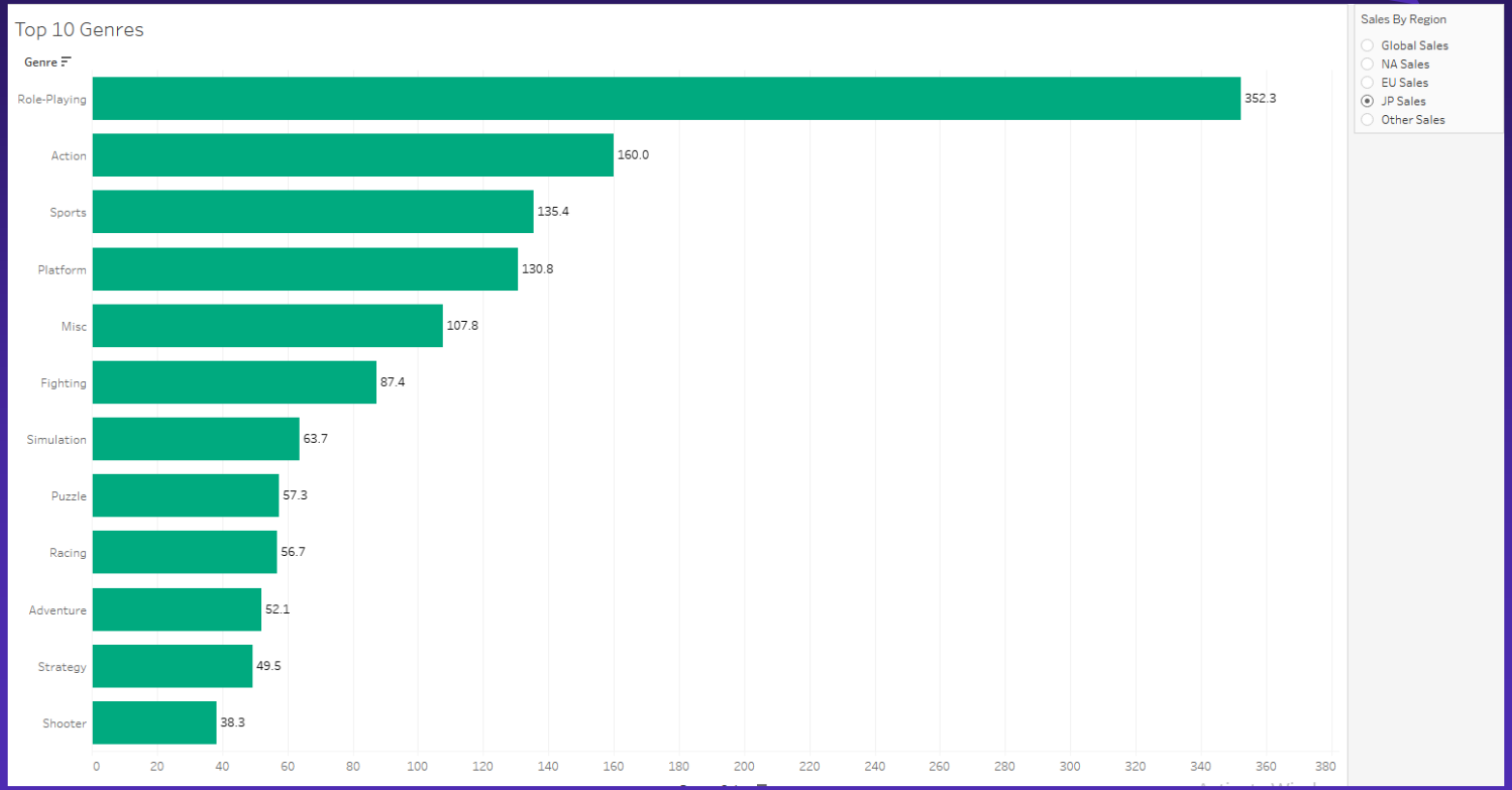
Data Analysis – Genre (EU)

- Top 10 Genres in Europe:



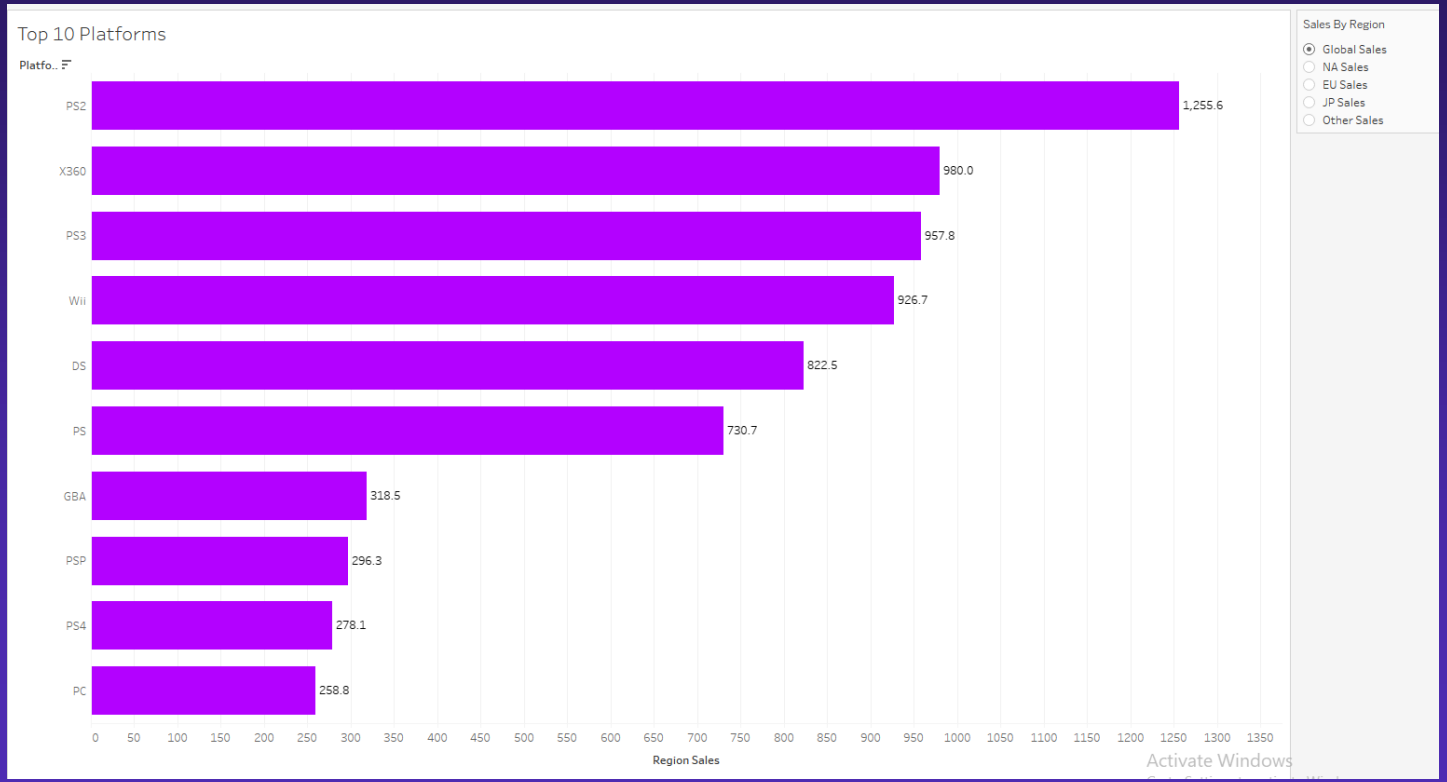
Data Analysis – Genre (Japan)

- Top 10 Genres in Japan:



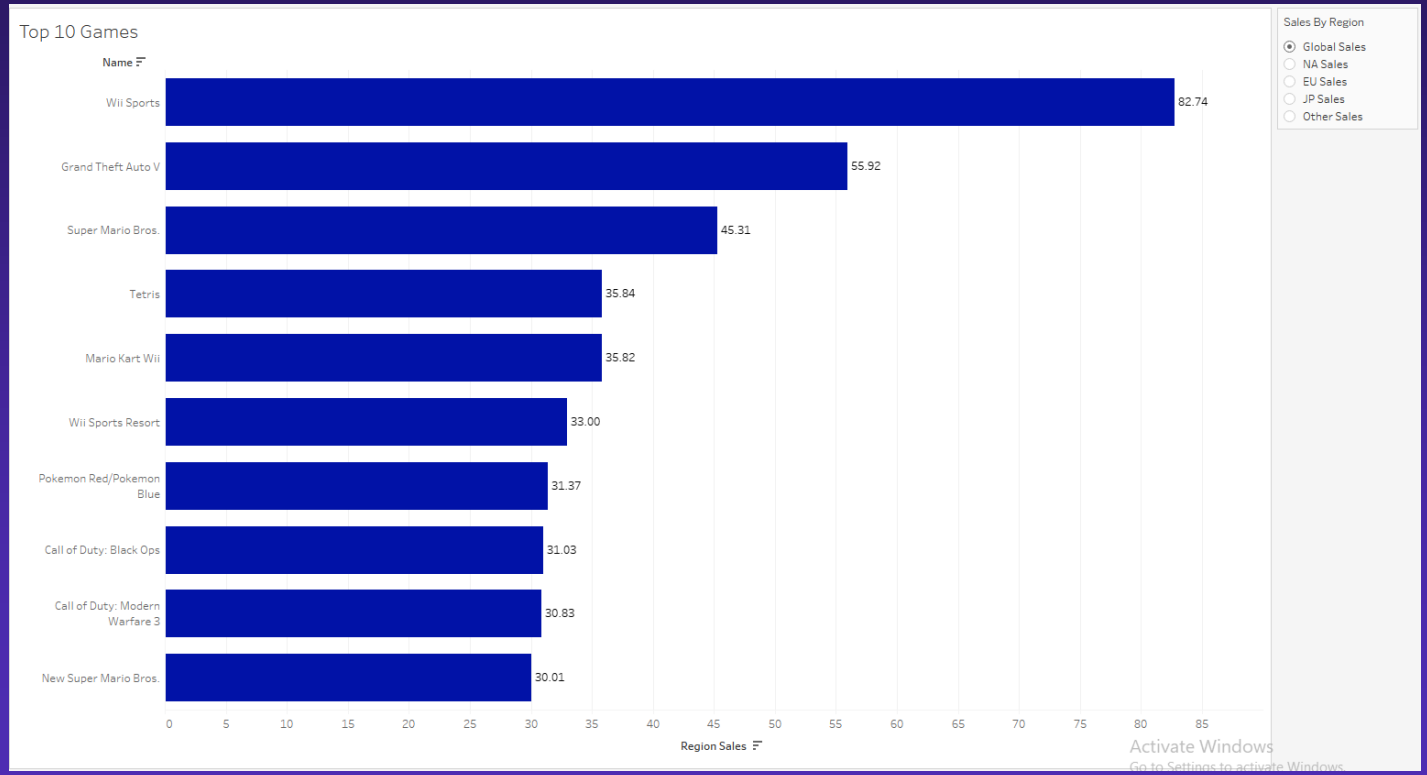
Data Analysis – Platform (Global)

- The top gaming platforms are as follows:



Data Analysis – Games

- The top games are as follows (adjustable by region):



Activate Windows
Go to Settings to activate Windows.

04

Insights & Recommendations

You can enter a subtitle here if you need it

Insights & Recommendations

- **GENRE:**
    - For both NA and EU, we notice a similar trend in the top few favorite gaming genres, such as Action, Sports, Shooter etc. However, for Japan, a distinct difference in the top genre is the Role-Playing category.
    - Using this data, video gaming companies in Japan could perhaps focus on producing games under the role-playing category (MMORPGs - Massive Multiplayer Online Role-Playing Games, or MMOs) such as Genshin Impact, Maplestory etc., and vice versa for the other regions.

Insights & Recommendations

- **PLATFORMS:**
    - The top gaming platform is the PS2, followed by XBOX360, PS3 & Wii. A possible reason for the PS2 being top could be that it has been around for a very long time, and many users have games which only the PS2 can play. Coupled with the fact that the new Playstation 5 is quite expensive, most users would prefer staying with their PS2.
    - A way to boost video game sales is to continue producing games which are playable and supported by the new PlayStations, as well as the older ones such as PS2.
    - However, if the aim is to boost the platform sales (such as PS5), the direction to look towards is to produce “Limited Edition” games that only the latest platforms can support. This might in turn boost the sales of the latest available gaming platforms.

Insights & Recommendations

- **SPECIFIC GAMES:**
    - The most popular game is the Wii Sports, followed by Grand Theft Auto V and Super Mario Bros.
    - The Wii Sports is a category of its own which contains the different sports games playable by Wii, which could be why it has the most sales out of the other games. However, gaming companies could leverage on this fact by producing and creating more Wii Sports games which are even more interactive physically, which is one of the unique attribute of Wii Sports that attracts potential video gamers.
    - On the other hand, Grand Theft Auto V was a highly anticipated game with its multiple successive franchise in the past. Also, Grand Theft Auto V was released on multiple platforms – which is probably the main reason for its popularity. Building on this concept, video game companies could do the same by releasing a single game on multiple platforms to boost their sales and revenue.

Conclusion

By transforming raw data into clean and useful data using platforms such as Excel & SQL, we can create charts and visualizations out of it using Tableau. With these visualizations, we can better understand the data presented to us, and make insightful recommendations to the relevant stakeholders, which could effectively boost their sales & revenue.

Thanks!

The End

CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon** and infographics & images by **Freepik**