

Imanol Hoyos Seiz
Data_science
Eda

¿Qué géneros y plataformas de videojuegos son las más exitosas?



Índice

- ¿Cómo se catalogan los videojuegos?
- Recopilación de los datos
- Análisis de los lanzamientos
- Análisis de las ventas
- Distribución por regiones
- Análisis de distribución por género y plataformas

¿Cómo se catalogan los videojuegos?

Antes de analizar los datos que se han recopilado para este EDA, primero es conveniente hablar un poco sobre cómo son catalogados los videojuegos según su género, lo cual es uno de los puntos que se han usado para analizar los datos.

Los géneros de videojuegos son unas categorías que nos sirven para clasificar a los mismos en función de distintos elementos, como por ejemplo sus gráficos o la jugabilidad de los mismos, siendo este último el factor más utilizado. Según este sistema existen una gran cantidad de géneros tales como el género de *Acción, Terror, First Person Shooter, Survival, Deportes...*, aunque esto en numerosas ocasiones puede ser un inconveniente, puesto que los límites de cada género no siempre están bien definidos y muchas veces queda a libre interpretación de los medios o incluso se suelen mezclar varios géneros distintos.

Recopilación de los datos

Para la realización de este EDA finalmente he decidido utilizar tres tablas encontradas en Kaggle y que estaban basadas en la web “www.vgchartz.com”. La primera más grande con 16.719 datos iniciales y que comprendía una gran variedad de géneros y plataformas, así como muchos de los videojuegos lanzados a lo largo de la historia, aunque desafortunadamente no estaba muy actualizada en los últimos años.

	Name	Platform	Year_of_Release	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales
0	Wii Sports	Wii	2006.0	Sports	Nintendo	41.36	28.96	3.77	8.45	82.53
1	Super Mario Bros.	NES	1985.0	Platform	Nintendo	29.08	3.58	6.81	0.77	40.24
2	Mario Kart Wii	Wii	2008.0	Racing	Nintendo	15.68	12.76	3.79	3.29	35.52
3	Wii Sports Resort	Wii	2009.0	Sports	Nintendo	15.61	10.93	3.28	2.95	32.77
4	Pokemon Red/Pokemon Blue	GB	1996.0	Role-Playing	Nintendo	11.27	8.89	10.22	1.00	31.37
...
16714	Samurai Warriors: Sanada Maru	PS3	2016.0	Action	Tecmo Koei	0.00	0.00	0.01	0.00	0.01
16715	LMA Manager 2007	X360	2006.0	Sports	Codemasters	0.00	0.01	0.00	0.00	0.01
16716	Haitaka no Psychedelica	PSV	2016.0	Adventure	Idea Factory	0.00	0.00	0.01	0.00	0.01
16717	Spirits & Spells	GBA	2003.0	Platform	Wanadoo	0.01	0.00	0.00	0.00	0.01
16718	Winning Post 8 2016	PSV	2016.0	Simulation	Tecmo Koei	0.00	0.00	0.01	0.00	0.01

16719 rows x 10 columns

Las otras dos tablas estaban más actualizadas aunque solo tomaban datos sobre las plataformas de PlayStation 4 (1.034 datos) y Xbox One (613 datos) respectivamente, por tanto las usé como apoyos para actualizar un poco la tabla base.

	Name	Platform	Year_of_Release	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales
0	Grand Theft Auto V	PS4	2014.0	Action	Rockstar Games	6.06	9.71	0.60	3.02	19.39
1	Call of Duty: Black Ops 3	PS4	2015.0	Shooter	Activision	6.18	6.05	0.41	2.44	15.09
2	Red Dead Redemption 2	PS4	2018.0	Action-Adventure	Rockstar Games	5.26	6.21	0.21	2.26	13.94
3	Call of Duty: WWII	PS4	2017.0	Shooter	Activision	4.67	6.21	0.40	2.12	13.40
4	FIFA 18	PS4	2017.0	Sports	EA Sports	1.27	8.64	0.15	1.73	11.80
...
1029	Fallen Legion: Flames of Rebellion	PS4	NaN	Role-Playing	NaN	0.00	0.00	0.00	0.00	0.00
1030	Radial G Racing Revolved	PS4	2017.0	Racing	Tammeka Games	0.00	0.00	0.00	0.00	0.00
1031	The Mummy Demastered	PS4	NaN	Action	NaN	0.00	0.00	0.00	0.00	0.00
1032	Project Nimbus: Code Mirai	PS4	NaN	Action	NaN	0.00	0.00	0.00	0.00	0.00
1033	Battle Chef Brigade	PS4	NaN	Action	NaN	0.00	0.00	0.00	0.00	0.00

1034 rows × 10 columns

	Name	Platform	Year_of_Release	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales
0	Grand Theft Auto V	XOne	2014.0	Action	Rockstar Games	4.70	3.25	0.01	0.76	8.72
1	Call of Duty: Black Ops 3	XOne	2015.0	Shooter	Activision	4.63	2.04	0.02	0.68	7.37
2	Call of Duty: WWII	XOne	2017.0	Shooter	Activision	3.75	1.91	0.00	0.57	6.23
3	Red Dead Redemption 2	XOne	2018.0	Action-Adventure	Rockstar Games	3.76	1.47	0.00	0.54	5.77
4	MineCraft	XOne	2014.0	Misc	Microsoft Studios	3.23	1.71	0.00	0.49	5.43
...
608	Biomutant	XOne	2018.0	Action	THQ Nordic	0.00	0.00	0.00	0.00	0.00
609	Biomutant	XOne	2019.0	Action	THQ Nordic	0.00	0.00	0.00	0.00	0.00
610	de Blob	XOne	2017.0	Platform	THQ Nordic	0.00	0.00	0.00	0.00	0.00
611	Outcast: Second Contact	XOne	2017.0	Adventure	Bigben Interactive	0.00	0.00	0.00	0.00	0.00
612	Code Vein	XOne	2019.0	Action	Bandai Namco Entertainment	0.00	0.00	0.00	0.00	0.00

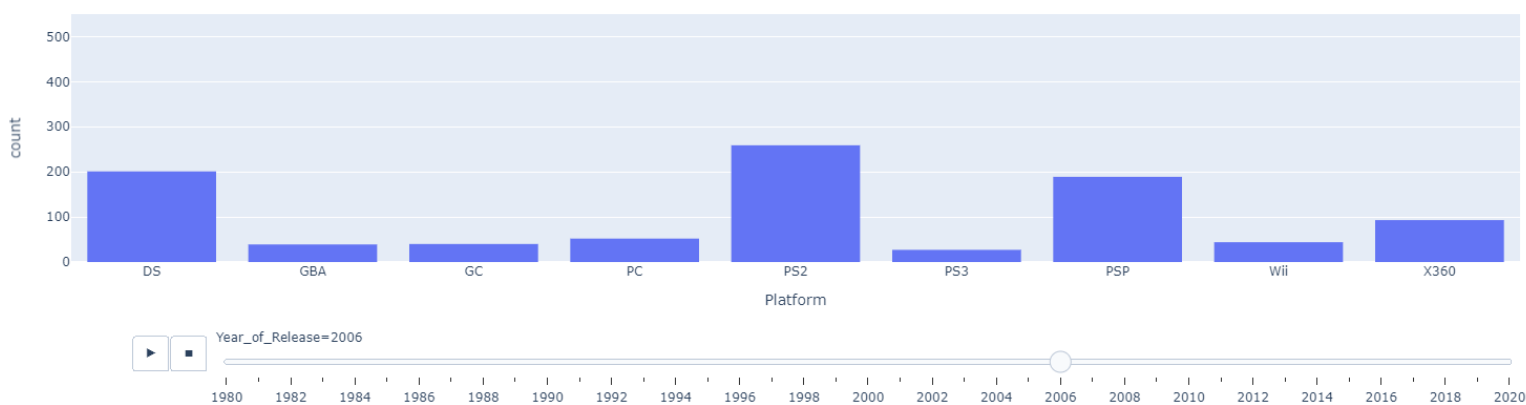
613 rows × 10 columns

Para limpiar la tabla final, eliminé los datos que se habían duplicado después de juntar las tres tablas, así como las filas con datos nulos, puesto que era un número muy reducido en comparación y eran principalmente juegos que no tenían en el dataset registrados su fecha de publicación o sus datos de ventas y no afectarían en exceso el resultado final. El resultado final fue una tabla con 17.703 entradas y fue la que usé para crear todas las gráficas de Plotly para la realización del EDA.

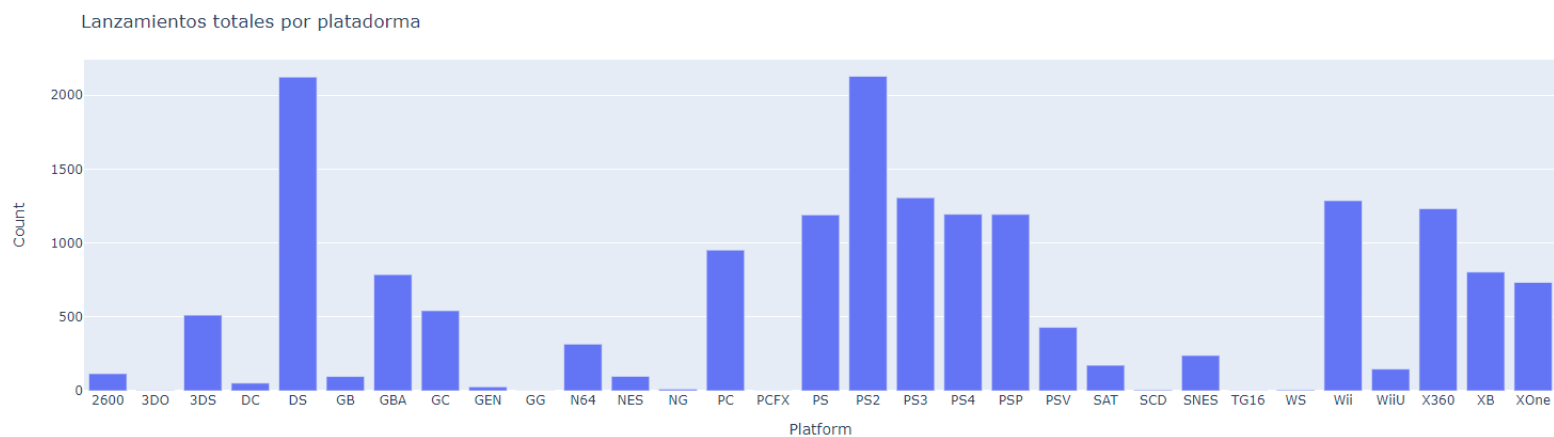
Análisis de los lanzamientos

En primer lugar quise analizar como se han ido distribuyendo los lanzamientos de videojuegos entre las diferentes plataformas y a lo largo de los años, para posteriormente poder compararlo con la distribución de las ventas.

Lanzamientos de videojuegos por año y plataforma

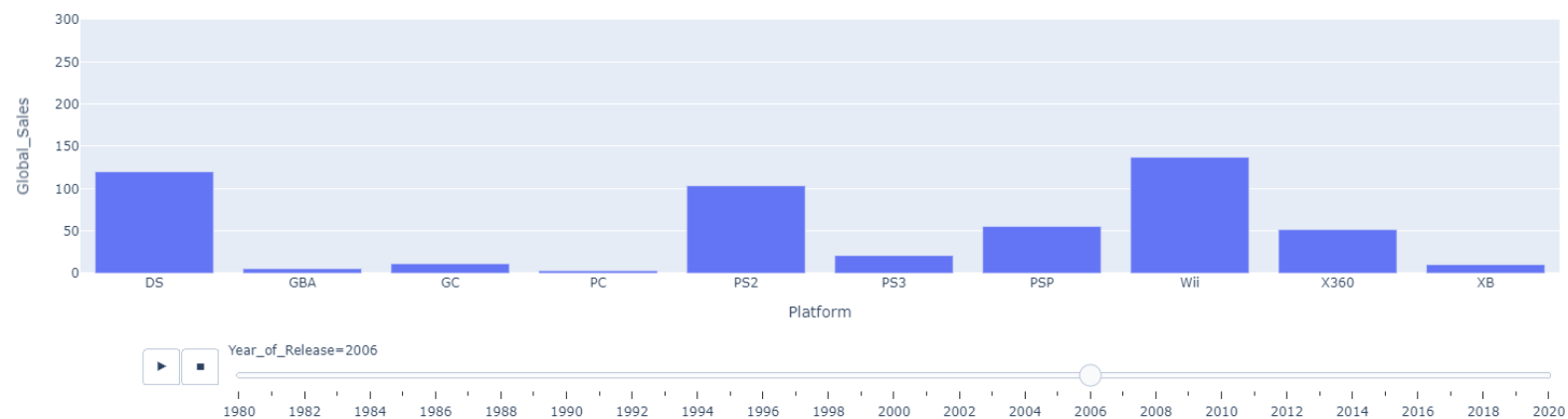


En esta línea temporal (referirse al archivo adjunto en el zip “Lanzamientos_Juegos_Plataforma.html”) se puede observar la evolución de los lanzamientos de las distintas plataformas a lo largo de los años, como al principio de crearse la industria con las primeras consolas la cantidad de videojuegos que salían cada año no era muy grande y esta fue creciendo a medida que la industria avanzaba, iban naciendo nuevas generaciones de consolas y este medio se iba instalando como una de las principales fuentes de ocio de las sociedad actual. Finalmente recogí en un gráfico la cantidad total de juegos lanzados por cada plataforma, pudiendo observar como la PlayStation 2 y la Nintendo DS se encuentran muy parejas, aunque supera la PlayStation 2 con 2.127 juegos (referir al archivo “Lanzamientos_Plataforma.html”).

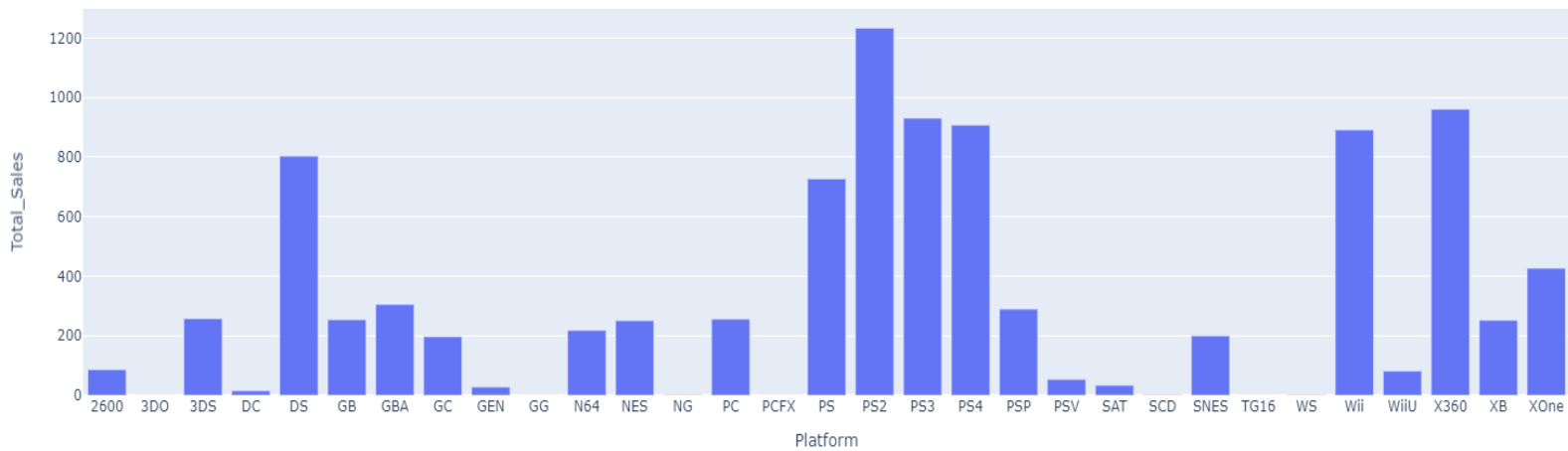


Análisis de las ventas

A continuación, repetí la operación con el historial de copias vendidas de cada videojuego a lo largo de los años en las diversas plataformas. Comparando la línea de tiempo de los lanzamientos con el de las copias vendidas (referir al archivo “Ventas_Año_Plataforma.html”), se puede observar que en general se mantiene la proporción de ventas a medida que avanzan los años en comparación con los lanzamientos analizados anteriormente.



Al igual que con los lanzamientos, en la siguiente gráfica agrupé el total de copias de juegos vendidos por cada plataforma (referir al archivo “Ventas_Plataforma.html”). De nuevo podemos observar que la ganadora fue la PlayStation 2 con 1.223,46 millones de juegos vendidos, aunque observando el conjunto global vemos en proporción con los lanzamientos de cada una, la Nintendo DS desciende considerablemente provocando que otras plataformas ganen un mayor protagonismo, como la PlayStation 3 o la Xbox 360.

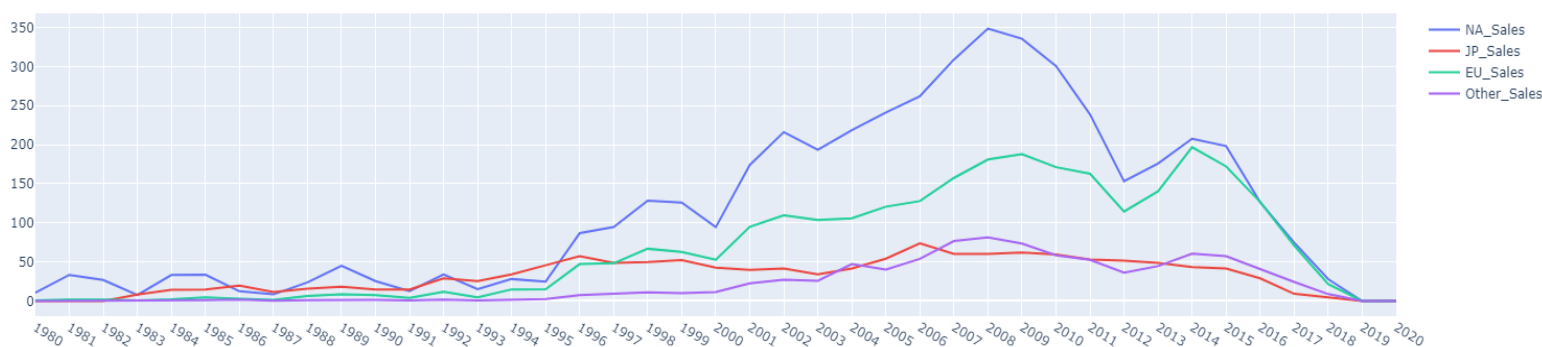


Por último en el caso de las ventas quise añadir un gráfico (referir al archivo “Juegos_Ventas.html”) para recoger los 20 juegos con el mayor número de copias vendidas a lo largo de la historia (aunque debido a que los datasets no estaban completamente actualizados, esta puede no ser del todo exacta), contando como ventas distintas las realizadas en diferentes plataformas para los juegos multiplataforma.



Distribución por Regiones

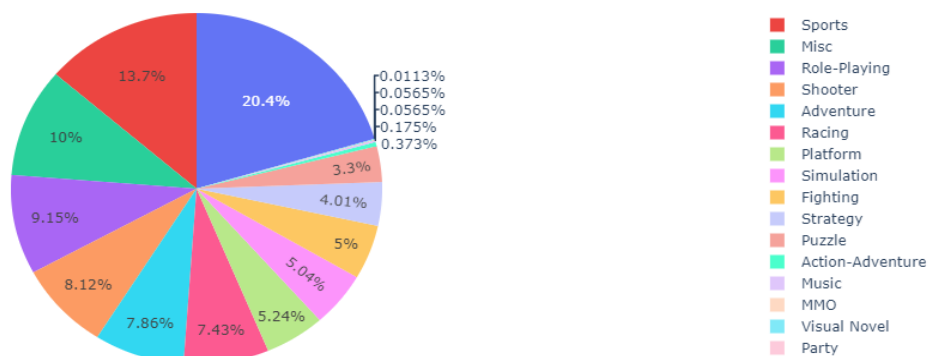
En el siguiente apartado del EDA se analiza la distribución de copias vendidas en diferentes regiones a lo largo de los años (referir al archivo “Ventas_Año_Región.html”), siendo estas *NA_Sales*, *JP_Sales*, *EU_Sales*, *Other_Sales*. De igual forma se observa el mismo patrón de crecimiento a lo largo que la industria se iba asentando gradualmente con los años aunque se puede observar una enorme diferencia con respecto a los dos principales mercados, Norte América y Europa, frente al resto de mercados, demostrando que estas suelen ser las poblaciones que más dirigen su ocio hacia el sector de los videojuegos de consola.



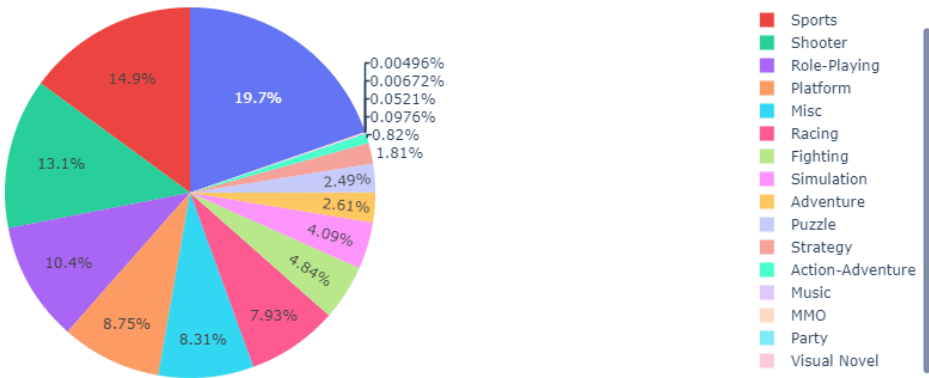
Análisis de distribución por Género y Plataforma

Como última parte del EDA hacemos al fin referencia a la pregunta inicial de “¿Qué géneros y Plataformas de videojuegos son las más exitosas?”. De igual forma que se analizó anteriormente los lanzamientos y las ventas de juegos por cada plataforma, a continuación se analizan en función de los diferentes géneros recogidos en estos datasets (referir a los archivos “Géneros_Lanzamientos.html” y “Géneros_Ventas.html” respectivamente). Se puede observar que los dos primeros puestos de *Acción* y *Sports* se mantienen con una proporción parecida en ambos casos, aunque en otros géneros existen diferencias como por el ejemplo con el género *Shooter*, que a pesar de tener un menor porcentaje de lanzamientos y situarse por debajo de otros géneros en este aspecto, si miramos a las ventas estas son bastante superiores, situándose en la tercera posición.

Lanzamientos por género



Géneros con más copias vendidas



Para terminar, en el último gráfico se analiza las ventas totales de videojuegos distribuidas en géneros y Plataformas simultáneamente para responder así a la pregunta inicial (referir al archivo “Géneros_Plataforma.html”). Teniendo en cuenta ambos aspectos, se puede observar que conjunto género-plataforma más exitoso se trata de *Action-PlayStation 3* (297,83 millones de copias vendidas), seguido de *Sport-Wii* (286,54 millones de copias vendidas). En cuanto al género vemos que se mantienen los dos primeros géneros analizados anteriormente, sin embargo en el aspecto de las plataformas esto no se mantiene, puesto que observando las ventas por plataforma vimos que las claras ganadoras eran la PlayStation 2 y la Xbox 360, mientras que analizando junto a los géneros, los resultados cambian a favor de la PlayStation 3 y la Wii, lo cual puede significar que aunque las dos anteriores tuvieran más ventas, los núcleos de estas se encuentran más distribuidos entre los diferentes géneros mientras que en las dos últimas estas ventas están más concentradas en unos géneros específicos.

