## **PROYECTO “Video Game Sales and Ratings”**

### **1. Introducción**

Los videojuegos son una de las formas más populares de entretenimiento a nivel mundial. El éxito de un juego depende de diversos aspectos como la plataforma en la que se juega, su género, las opiniones de los críticos y usuarios, así como las ventas en distintas regiones. A veces no es fácil ver las conexiones entre los datos, pero al analizarlos es posible encontrar patrones que puedan explicar por qué algunos videojuegos tienen más éxito que otros en el mercado. Este proyecto analiza datos mundiales de videojuegos para estudiar cómo la plataforma, el género y las calificaciones influyen en las ventas a nivel global y regional.

**2. Objetivo general**

Procesar un conjunto de datos de videojuegos para explorar la relación entre las ventas globales y regionales, las plataformas, los géneros, las puntuaciones de críticos y usuarios, con el fin de identificar patrones que puedan influir en el éxito comercial de los videojuegos.

### **3. Objetivos específicos**

* **Limpiar y organizar** la base de datos obtenida para garantizar que los datos estén listos para su análisis.
* **Analizar la relación** entre las ventas globales y las ventas por región (Norteamérica, Europa, Japón y otras) para determinar si alguna región domina el mercado.
* **Comparar el éxito** de diferentes plataformas (PS4, Xbox, PC, etc.) en las ventas globales y regionales.
* **Evaluar la influencia** de las puntuaciones de críticos y usuarios en las ventas totales de los videojuegos.
* **Explorar las tendencias** de ventas según el género del videojuego identificando los géneros más exitosos en cada región.

### **4. Planteamiento del problema**

La industria de los videojuegos enfrenta un desafío constante: identificar los factores clave que determinan el éxito comercial de un título. Si bien elementos como la plataforma y el género de un juego son considerados determinantes, aspectos adicionales como las calificaciones de críticos, las preferencias de diferentes regiones y el momento del lanzamiento también pueden desempeñar un papel crucial. En un mercado en constante crecimiento, se espera que las ventas de un videojuego puedan ser predichas en función de las tendencias y las preferencias de los consumidores. Sin embargo, existe una desconexión frecuente entre las calificaciones y las ventas: un juego altamente valorado no siempre se traduce en grandes ventas, y los títulos más vendidos no necesariamente obtienen las mejores críticas. Este proyecto busca analizar un conjunto de datos de videojuegos con el fin de descubrir qué características tienen mayor influencia en las ventas globales y regionales, y cómo estas variables están interrelacionadas.

### **5. Preguntas de investigación**

* ¿Cuántos videojuegos tienen información incompleta o nula en términos de ventas o puntuaciones?
* ¿Cuál es el rango de ventas en las diferentes regiones (Norteamérica, Europa)?
* ¿Qué plataformas generan mayores ventas globales?
* ¿Qué géneros de videojuegos son los que mayor se venden?
* ¿Cuál es la distribución de los videojuegos por género y plataforma?
* ¿Qué relación existe entre las puntuaciones de críticos y usuarios y las ventas globales?
* ¿Existen diferencias significativas en las ventas según el año de lanzamiento de los videojuegos?
* ¿Hay juegos con buenas calificaciones que venden poco? ¿Y juegos mal calificados que venden mucho?

### **6. Posible solución**

La solución consistirá en implementar un proceso de limpieza y procesamiento de datos utilizando Python y las bibliotecas de **pandas** y **numpy**. Se abordarán tareas como la eliminación de valores nulos, la normalización de las puntuaciones, y el cálculo de las ventas totales por región. Luego, se realizarán análisis estadísticos para identificar patrones en las ventas según las características de los videojuegos. Visualizaciones como gráficos de barras y de líneas serán empleadas para facilitar la interpretación de los resultados y responder las preguntas planteadas.

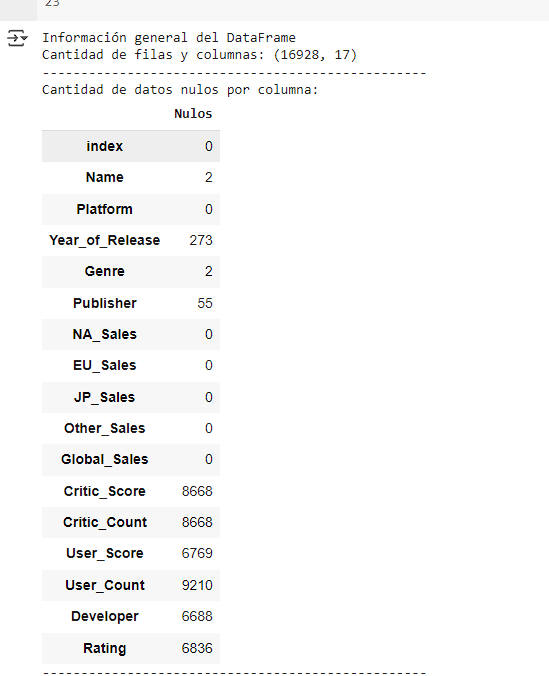
### **7. Pasos y resultados**

1. **¿Cuántos videojuegos tienen información incompleta o nula en términos de ventas o puntuaciones?**

La función isnull() busca los valores nulos, la aplicamos en un principio para la limpieza de datos, nos muestra la cantidad de datos nulos en cada columna.



**Resultados**

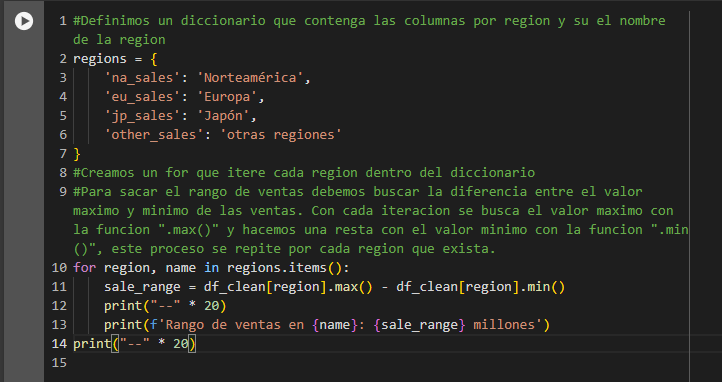
****

El resultado de la ejecución fueron diversos, en ventas no tenemos valores nulos por tanto no se modificará, en el caso de puntuaciones obtenemos bastantes valores NaNs nuestra solución fue llenar estos datos nulos con fillna() reemplazando los valores de las puntuaciones con valores generados aleatoriamente en base al promedio de los datos de la columna correspondiente.

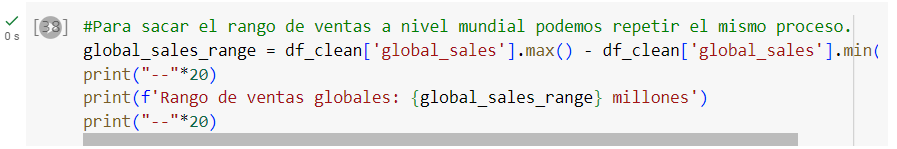
1. **¿Cuál es el rango de ventas en las diferentes regiones (Norteamérica, Europa, Japón)?**

El código calcula el rango de ventas de videojuegos en distintas regiones (Norteamérica, Europa, Japón y otras) mediante la diferencia entre las ventas máximas y mínimas registradas en cada área. Se utilizó un diccionario que asocia las columnas de ventas con nombres reales, se itera sobre cada región, se calcula el rango de ventas y se manda a imprimir..

Por Región:



Global:

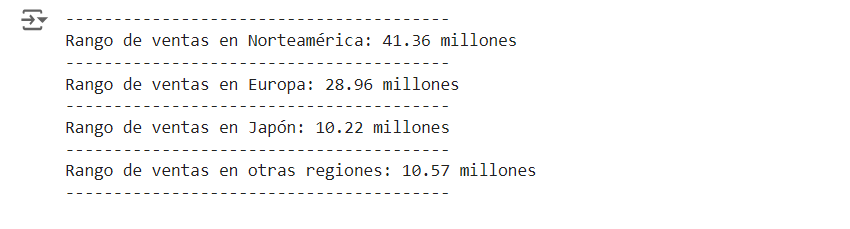


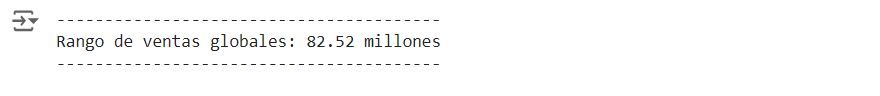
Se utilizó la función `.max()` para encontrar el valor máximo de ventas y la función `.min()` para el valor mínimo.

La diferencia entre estos dos valores se almacenó en variables específicas para cada región.

**Resultados**

Por Región:

****Global:

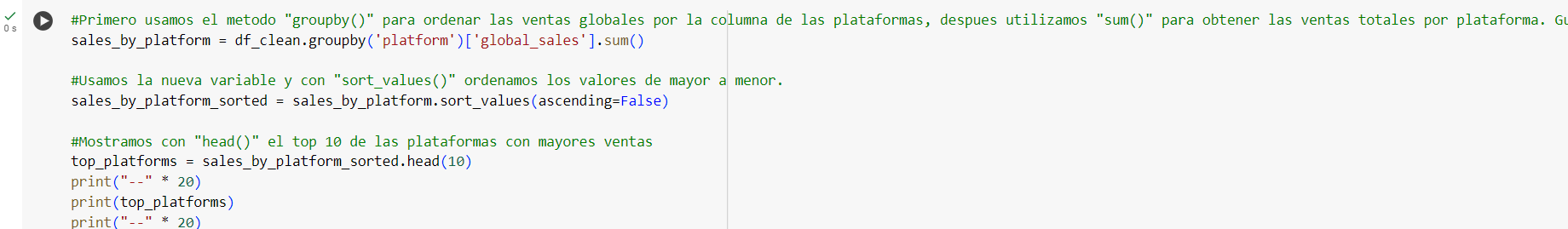


Estos resultados indican que hay variaciones significativas en las ventas de videojuegos entre las diferentes regiones. La cobertura de ventas en Norteamérica y Europa es particularmente alta en comparación con otras regiones, lo que indica que estas son regiones clave en el mercado de los videojuegos.

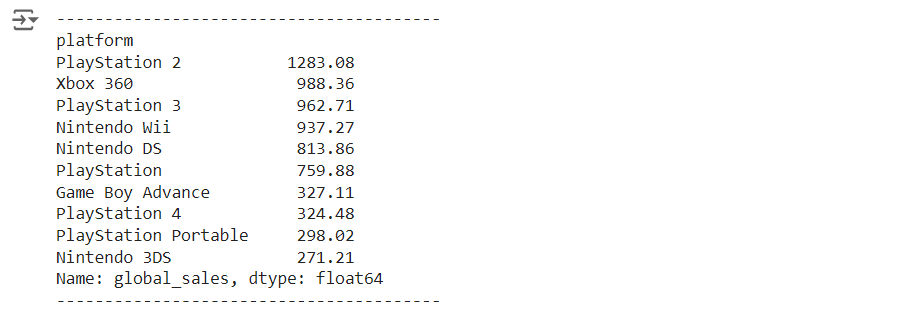
Las ventas globales de 82,52 millones demuestran el potencial y la diversidad del mercado global de videojuegos, aunque también indica que las ventas pueden variar significativamente entre los diferentes títulos de juegos.

1. **¿Qué plataformas generan mayores ventas globales?**

Para agrupar los datos de ventas por plataforma se utilizó groupby(), además de sum() para obtener el total de ventas globales, posteriormente sort\_values(ascending=False) para ordenar las plataformas de mayor a menor en función a lo anterior.



**Resultados**

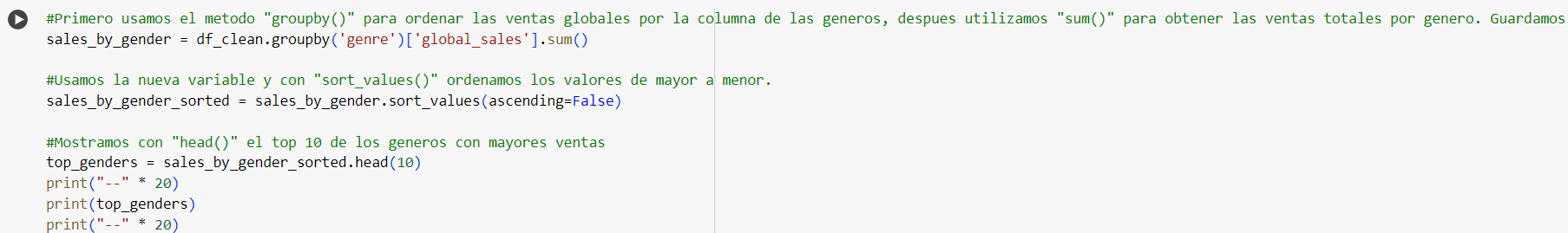
****

PlayStation 2 es la consola con mayores ingresos globales en la historia de los videojuegos, seguida de Xbox 360 y PlayStation 3.

Las consolas de Nintendo como Nintendo Wii y Nintendo DS también se encuentran entre las consolas más vendidas, lo que demuestra su fuerte presencia en el mercado.

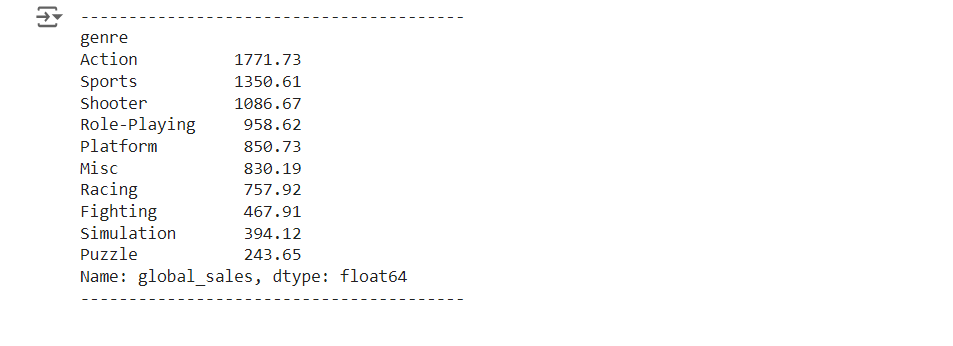
1. **¿Qué géneros de videojuegos son los que mayor se venden?**

Se realizó un análisis de las ventas globales totales de cada género de videojuego. Se agruparon los datos por género y luego se sumaron las ventas globales para cada uno de ellos.



**Resultados**

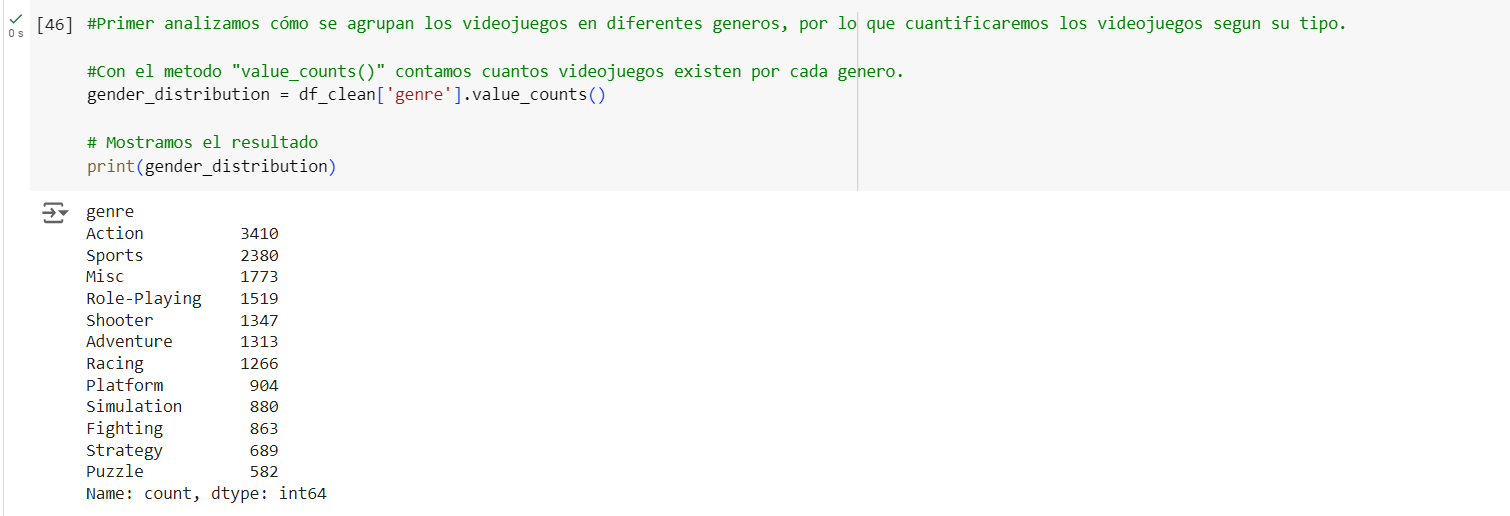
Los géneros más vendidos son Action (1,771.73 millones), Sports (1,350.61 millones) y Shooter (1,086.67 millones), destacando la popularidad de los juegos dinámicos y competitivos a nivel mundial, además ofrecen experiencias intensas y variadas destacan entre su popularidad.



1. **¿Cuál es la distribución de los videojuegos por género y plataforma?**

Se realizó un análisis de cómo se agrupan los videojuegos según su género y plataforma.

Iniciamos cuantificando los videojuegos según su tipo.



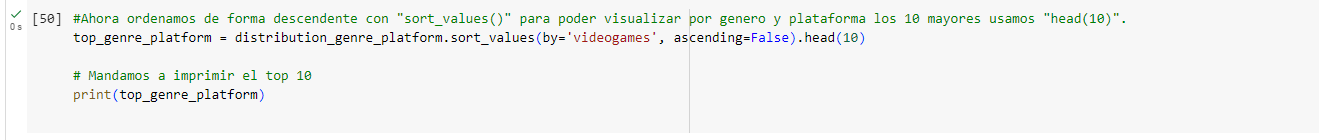
Después por plataforma.



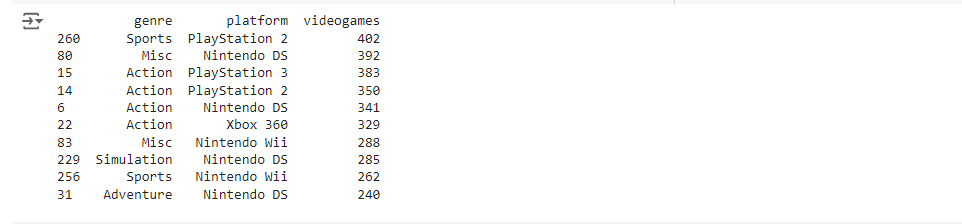
Posteriormente ambos (género y plataforma)



Para finalmente ordenarlos con el top 10, los más relevantes.



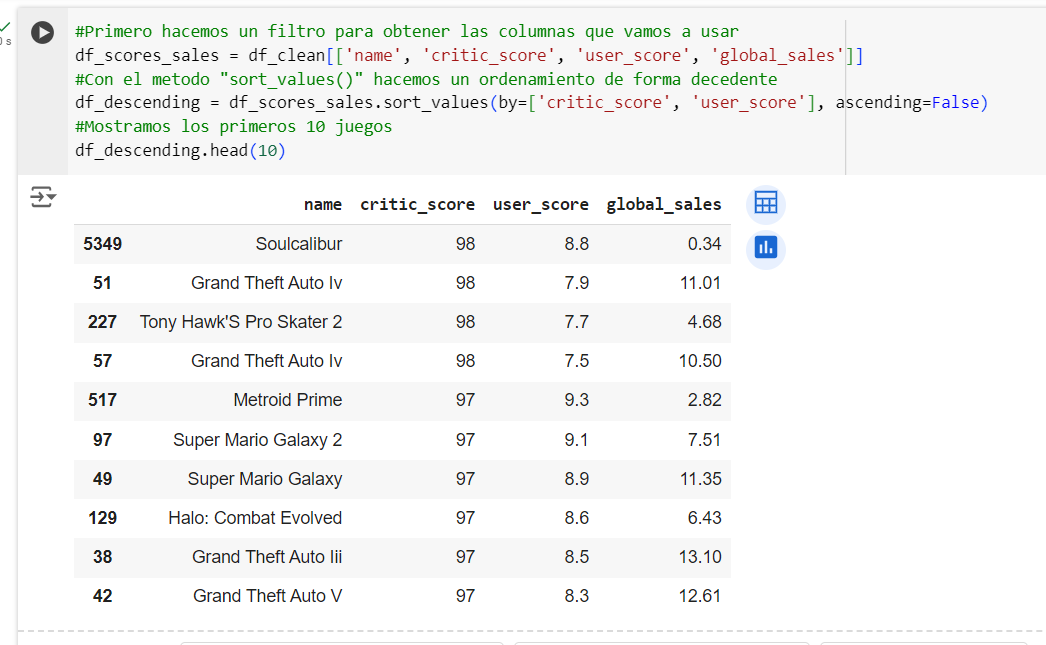
**Resultados**

****

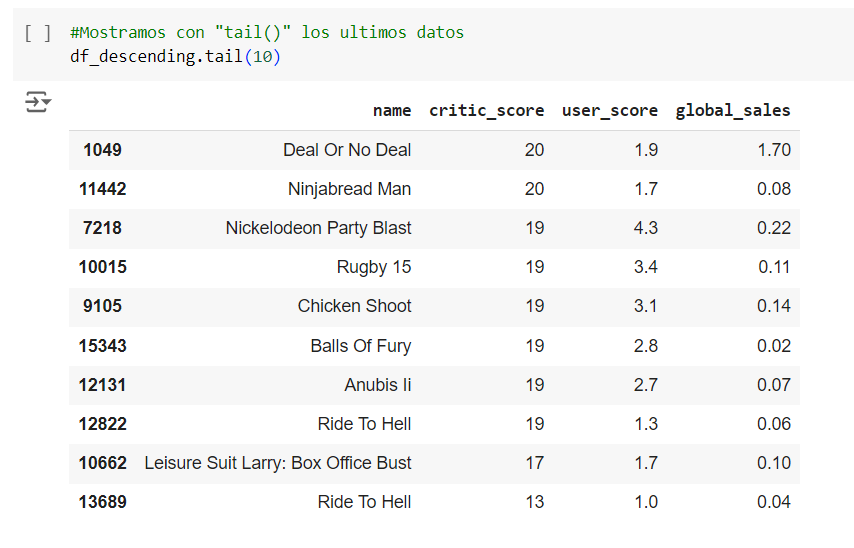
Los resultados muestran que Action (3,410), Sports (2,380), y Misc (1,773) son los géneros con más videojuegos. En términos de plataformas, las más populares son PlayStation 2 (2,188), Nintendo DS (2,164), y PlayStation 3 (1,359). Al combinar género y plataforma, se observa que Sports en PlayStation 2 tiene la mayor cantidad de videojuegos (402), seguido de Misc en Nintendo DS (392) y Action en PlayStation 3 (383).

1. **¿Qué relación existe entre las puntuaciones de críticos y usuarios y las ventas globales? ¿Hay juegos con buenas calificaciones que venden poco? ¿Y juegos mal calificados que venden mucho?**

Primero se filtraron las columnas relevantes para el análisis, seleccionando el nombre del juego, las puntuaciones de críticos y usuarios, y las ventas globales. Luego, los juegos se ordenan en forma descendente según las puntuaciones más altas usando la función sort\_values(). Finalmente, se imprimen los 10 juegos mejor calificados para observar si existe una relación entre las altas puntuaciones y las mayores ventas.



Se utiliza la función \*\*tail()\*\* para mostrar los últimos 10 juegos con las peores calificaciones. A pesar de las puntuaciones muy bajas de críticos y usuarios, estos juegos registran algunas ventas, aunque considerablemente menores.



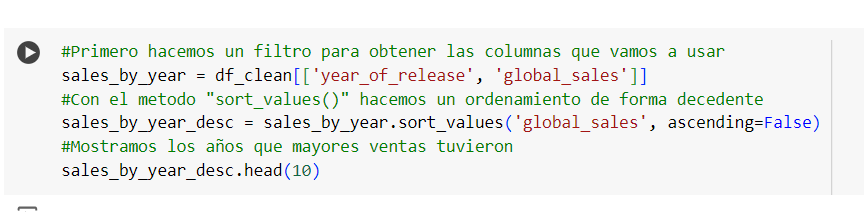
**Resultados**

Juegos como Grand Theft Auto V y Super Mario Galaxy tienen calificaciones excelentes de críticos (97) y usuarios, con ventas que superan los 10 millones de copias, mostrando una relación positiva en estos casos.

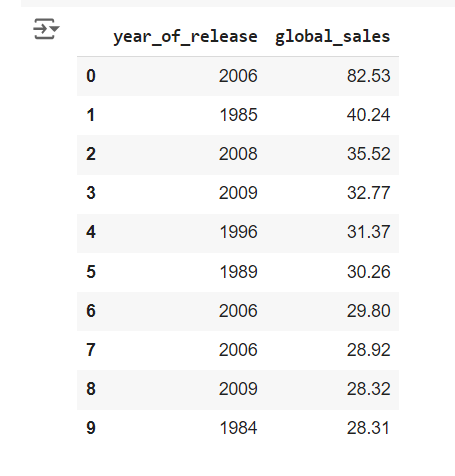
Mientras que juegos como Ride to Hell y Ninjabread Man tienen puntuaciones extremadamente bajas (entre 13 y 20), y las ventas son mínimas, mostrando que en estos casos las malas críticas sí afectan las ventas.

1. **¿Hay juegos con buenas calificaciones que venden poco? ¿Y juegos mal calificados que venden mucho?**

Se filtran las columnas relevantes para el análisis, seleccionando el año de lanzamiento y las ventas globales. Luego, se ordenan los datos de forma descendente utilizando sort\_values(). Finalmente, se muestran los 10 años con mayores ventas.



**Resultados**



Los datos muestran que algunos años, como 2006 y 1985, tuvieron ventas significativamente más altas que otros.

### **8. Consideraciones futuras**

Una vez que los datos estén procesados y limpios, será posible realizar un análisis más profundo. Por ejemplo, se podrían explorar correlaciones avanzadas entre las puntuaciones de los críticos y las ventas globales. Además, este proyecto podría servir como base para crear modelos predictivos que anticipen el éxito comercial de futuros videojuegos basándose en sus características clave. También podría ser interesante analizar el impacto de factores externos como campañas de marketing o lanzamientos simultáneos de consolas.

### **9. Conclusión**

En este análisis, hemos examinado diversos aspectos clave del mercado de videojuegos, revelando patrones sobre las plataformas, géneros y su relación con las ventas y calificaciones.

El procesamiento de datos es esencial para comprender las tendencias en la industria de los videojuegos y su relación con las ventas globales. En este proyecto, nos enfocamos en limpiar y analizar un conjunto de datos globales de videojuegos, permitiendo futuras investigaciones sobre las relaciones entre plataformas, géneros, puntuaciones y ventas. Con una base de datos procesada, será posible avanzar hacia análisis más complejos y construir modelos predictivos que apoyen la toma de decisiones estratégicas en el desarrollo y publicación de videojuegos.

1. Plataformas con mayores ventas globales: Las consolas PlayStation 2, Xbox 360 y PlayStation 3 dominan el mercado, demostrando que la calidad, el acceso a juegos icónicos y una sólida base de jugadores son fundamentales para el éxito comercial. Estas plataformas fueron líderes durante generaciones cruciales en la evolución de los videojuegos.

2. Géneros más vendidos: Los videojuegos de Acción, Deportes y Shooter generan las ventas más altas a nivel global. Esto refleja una clara preferencia por juegos con alta adrenalina, interacción directa y experiencias inmersivas, atrayendo tanto a jugadores casuales como dedicados.

3. Distribución por género y plataforma: Al analizar la distribución de títulos por género y plataforma, se observa que la PlayStation 2 y la Nintendo DS ofrecen una amplia variedad de juegos, especialmente en los géneros de Acción y Deportes. Esto subraya la importancia de contar con un catálogo diverso para mantener la competitividad y satisfacer una amplia gama de preferencias.

4. Relación entre puntuaciones y ventas: Aunque existe una correlación positiva entre altas calificaciones (de críticos y usuarios) y ventas en títulos como Grand Theft Auto V y Super Mario Galaxy, también se encontraron ejemplos de juegos con bajas puntuaciones (como Ride to Hell) que lograron ventas significativas. Esto sugiere que factores como la nostalgia, el marketing y la franquicia a la que pertenecen pueden impulsar las ventas, incluso ante críticas negativas.

En resumen, nuestro análisis no solo responde a preguntas clave sobre los factores que impulsan las ventas, sino también a cómo las preferencias de los jugadores, la atracción de los géneros y las estrategias de las plataformas se combinan para crear fenómenos globales en la industria del videojuego.

### **10. Entregables**

* Código ejecutable en Google Colab o Jupyter Notebook que realice el procesamiento y análisis de datos.
* Documento en Google Docs o Word que describa los pasos y resultados del procesamiento y análisis de datos.
* Presentación en PowerPoint o Google Slides con un resumen de las técnicas aplicadas y los resultados obtenidos en el análisis de los datos.