



# Ventas de videojuegos

Análisis de datos de ventas de más de 16.500  
juegos.

El mercado de los videojuegos ha tenido un aumento exponencial en las últimas décadas por lo cual el análisis de los datos relacionados implica un enorme potencial para entender el comportamiento de los consumidores. Incluso para la industria relacionada a los e-sports, la cual resulta muy importante en la actualidad.

Nuestro principal objetivo es analizar la historia de los videojuegos a través del conjunto de datos con el fin de comprender las causas por las cuales un videojuego resultó más exitoso que otro y tomar como base estos conocimientos y que resulten útiles para futuras tomas de decisión.



## Objetivos planteados

**Agrupación:** A partir de un algoritmo de agrupación, se necesita generar n cantidad de grupos según las características del dataset para luego caracterizar cada grupo en base a sus atributos.

**Predicción:** ¿Cuáles serán las ventas globales de un videojuego específico en función de sus características, como el nombre, la plataforma, el año de lanzamiento, el género y el editor?



## Resto de preguntas:

- ¿Cuál es el género más vendido en base a las ventas de una región según la empresa de videojuegos?
- ¿Cuál fue la empresa de videojuegos que obtuvo mayor cantidad de ganancias a nivel mundial entre 2000 y 2010?
- ¿Cuál es el juego más vendido de cada una de las principales empresas?
- ¿Cuál es el juego más vendido para cada tipo de género ?
- ¿Cuál fue el género de videojuego más consumido de cada empresa?
- ¿Cuáles son los juegos más vendidos en cada región geográfica (por ejemplo, Norteamérica, Europa, Japón)?
- ¿Cuál es la tendencia de ventas de videojuegos por género en Norteamérica comparado con Japón a lo largo de los años?
- ¿En qué año se registraron las ventas globales más altas en la historia de los videojuegos, y qué eventos o lanzamientos pueden haber contribuido a este logro?
- ¿Cuál es la empresa de videojuegos con el mayor número de títulos lanzados desde 2000 a 2016?
- ¿Cuál es la empresa que obtuvo mayor cantidad de ganancias en la región de Norte América para videojuegos tipo "Shooter"?
- Luego de que salga a venta la PS4, en base a las ganancias globales¿Tuvo algún impacto negativo la plataforma PS3?
- Tomando de referencia las ventas en Japón, ¿qué plataforma tuvo un impacto positivo en los últimos 6 años (2010-2016)?
- Se hará la agrupación de sagas de juegos y en base a esas agrupaciones de distintas sagas se podrán realizar distintas preguntas en base a ¿cuál es la saga que tuvo mayor éxito? ¿en qué región? dentro de una saga, ¿qué juego fué el que tuvo mayor ventas vendidas?



# Etapa 1: Preparación de datos

Preparar los datos para hacerlos accesibles a cualquier técnica estadística



# Motivos de la elección del dataset

- *Relevancia en la industria de los videojuegos*: puede proporcionar información valiosa sobre tendencias de ventas, géneros populares y plataformas exitosas.
- *Tomar decisiones empresariales*: en una empresa relacionada con los videojuegos o en el campo del entretenimiento digital, el EDA de este dataset podría ayudar a tomar decisiones comerciales informadas. Por ejemplo, se podría identificar qué géneros de juegos tienen un historial exitoso de ventas en ciertas regiones geográficas.
- *Explorar tendencias a lo largo del tiempo*: dado que el dataset incluye información sobre el año de lanzamiento de los juegos, se puede realizar un análisis de tendencias a lo largo del tiempo.
- *Comprender el mercado global*: el dataset incluye ventas desglosadas por regiones geográficas (como Norteamérica, Europa, Japón, etc.). Esto permitiría comprender mejor las diferencias en las preferencias de los jugadores en diferentes partes del mundo y cómo influyen en las ventas globales.



# Acerca del conjunto de datos

El conjunto de datos está conformado por una lista de videojuegos con ventas superiores a 100.000 copias, posee 16598 filas y 11 columnas. Fue generado por un scrapping del sitio web vgchartz, cuyo propósito es el monitoreo de ventas de videojuegos que ofrece ventas semanales de videoconsolas y juegos por región. Los campos que incluye el dataset son:

- **Rank:** Ranking de ventas generales
- **Name:** Nombre del juego.
- **Platform:** Plataforma de lanzamiento del juego (es decir, PC, PS4, etc.)
- **Year:** Año de lanzamiento del juego.
- **Genre:** Género del juego
- **Publisher:** Editor del juego
- **NA\_Sales:** Ventas en Norteamérica (en millones)
- **EU\_Sales:** Ventas en Europa (en millones)
- **JP\_Sales:** Ventas en Japón (en millones)
- **Other\_Sales:** Ventas en el resto del mundo (en millones)
- **Global\_Sales:** Ventas totales a nivel mundial (en millones).

# Análisis Exploratorio de Datos

Como punto de partida se procedió a realizar un informe completo de análisis exploratorio de datos (EDA) a partir del dataset elegido. El mismo se realizó utilizando las herramientas recomendadas por la cátedra (Pandas junto con Profiling). Esta técnica permitió obtener una vista general rápida y detallada del conjunto de datos, logrando así una comprensión exhaustiva de la estructura, calidad y contenido de los datos antes de realizar un análisis más profundo. En resumen, el informe generado proporciona una visión general detallada de los datos, lo que facilita la comprensión y la detección de problemas en el conjunto de datos.



# Descripción general del dataset utilizando YData Profiling

Overview

Alerts 15

Reproduction

## Dataset statistics

Number of variables	11
Number of observations	16598
Missing cells	563
Missing cells (%)	0.3%
Duplicate rows	0
Duplicate rows (%)	0.0%
Total size in memory	1.4 MiB
Average record size in memory	88.0 B

## Variable types

Numeric	7
Unsupported	2
Categorical	1
Text	1



# Data Wrangling

El data wrangling se llevó a cabo con el objetivo de preparar los datos de manera que sean coherentes, limpios y adecuados para realizar un análisis significativo.

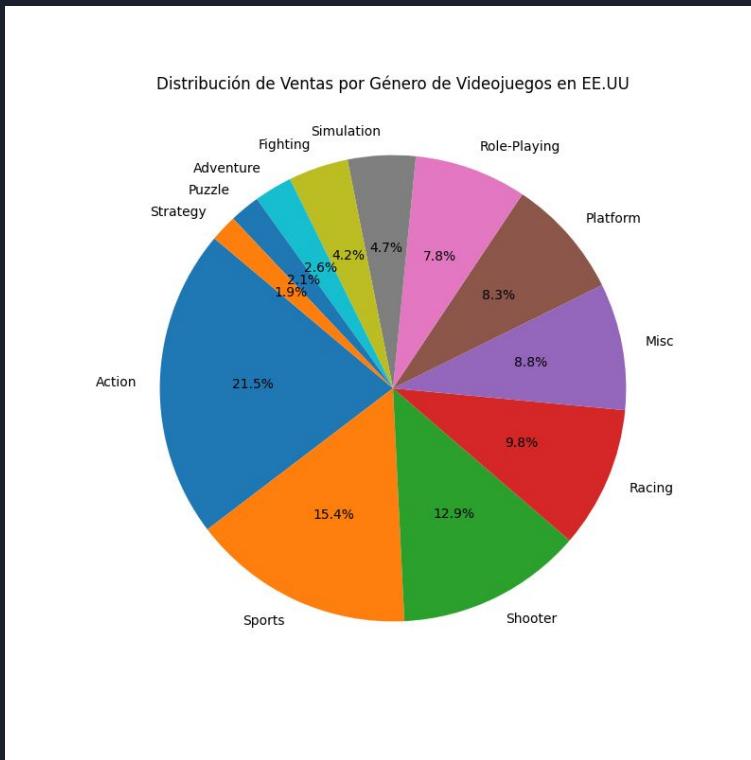
- Se optó por la eliminación de las celdas faltantes distribuidas en todo el dataset, las cuales eran 563 de 16598.
- Se corrigieron los tipos de datos que tenían las columnas ‘Name’ y ‘Publisher’ a texto y categórico debido a que según el reporte, ambas columnas tenían tipos de datos incompatibles.
- Se corrigió el tipo de dato ‘float64’ que contenía la columna ‘Year’ y se lo convirtió al tipo de dato ‘int32’.
- El dataset antes de ser limpiado y filtrado contaba con 203 filas con valores “unknown” pero luego de la eliminación de las celdas faltantes (NaN) quedaron 100. Por lo tanto, se optó en que estas filas permanezcan como “unknown” ya que los mismas no representan una cantidad significativa.



## Etapa 2: Análisis estadístico

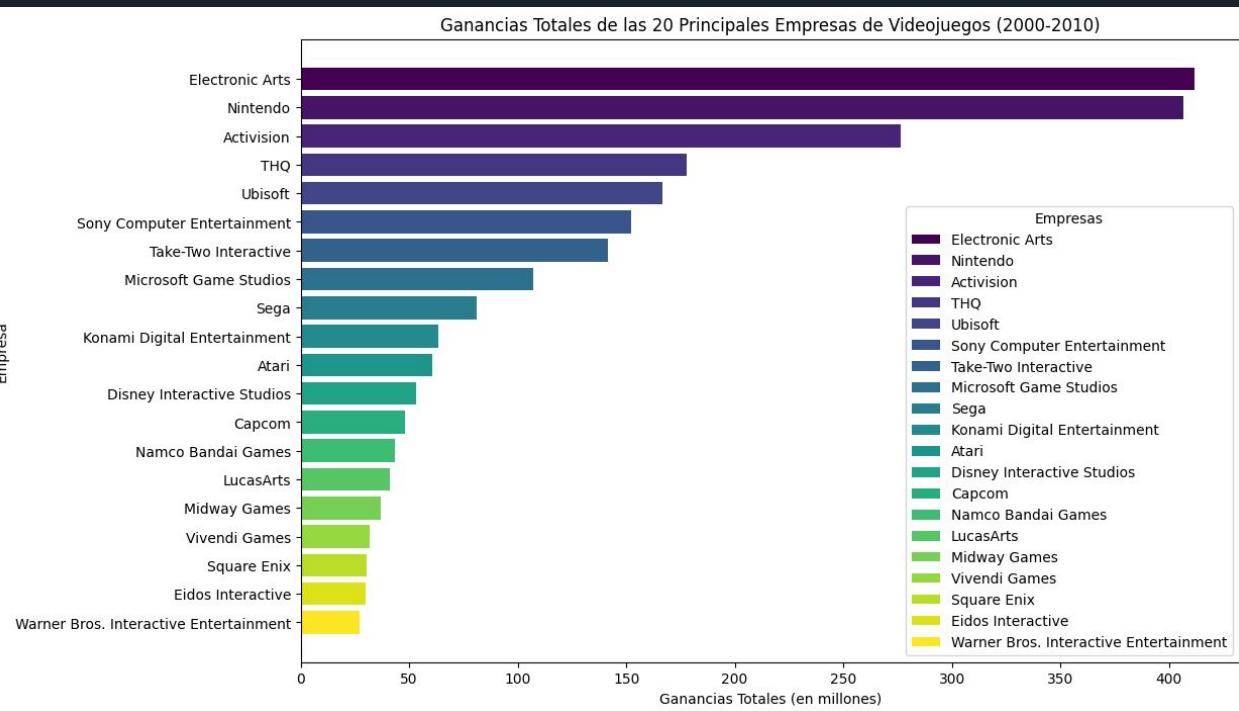
Realizar un examen gráfico de la naturaleza de las variables individuales a analizar y un análisis estadístico que permita cuantificar algunos aspectos gráficos de los datos

# 1- ¿Cuál es el género más vendido en base a las ventas de una región según la empresa de videojuegos?



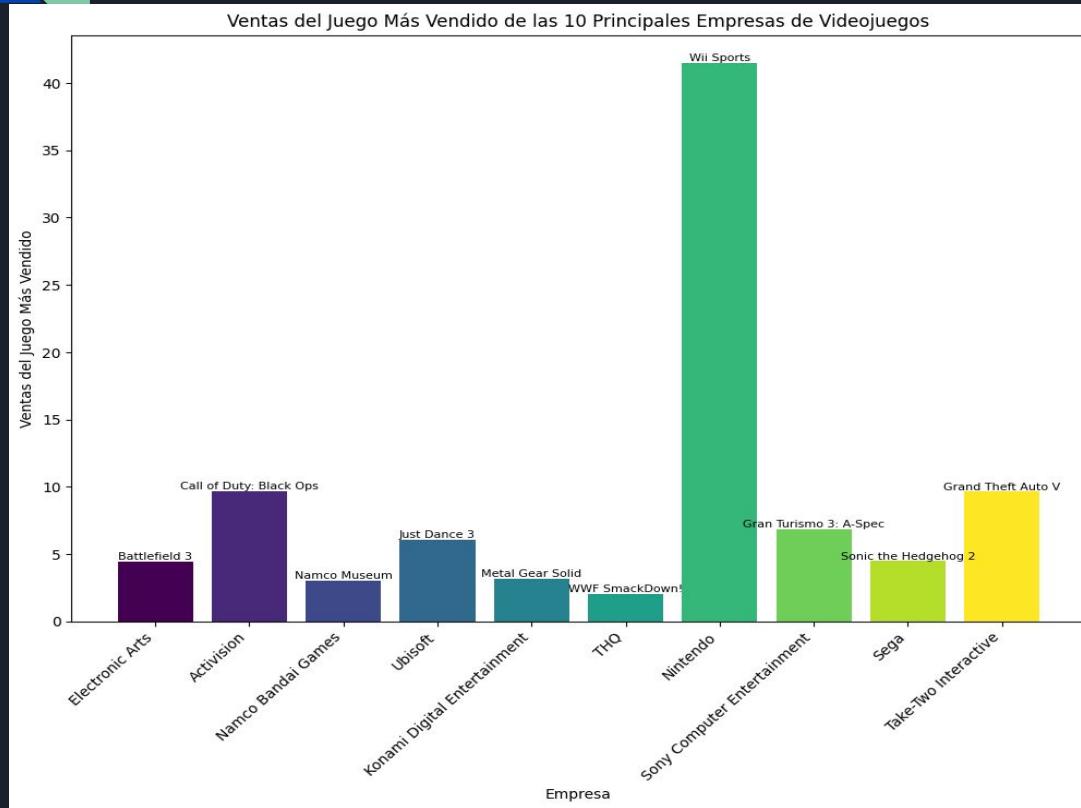
A partir del gráfico de pastel, se observa claramente que el género más vendido para la empresa Nintendo en Estados Unidos es el "Action" (Acción), seguido de cerca por el género "Sports" (Deportes). Estos dos géneros dominan significativamente las ventas, mientras que los demás géneros tienen una presencia mucho menor en comparación. Esta información es crucial para entender las preferencias de los jugadores en **Estados Unidos** en relación con los productos ofrecidos por **Nintendo**, lo que puede ser útil para futuras estrategias de marketing y desarrollo de juegos.

## 2- ¿Cuál fue la empresa de videojuegos que obtuvo mayor cantidad de ganancias a nivel mundial entre 2000 y 2010?



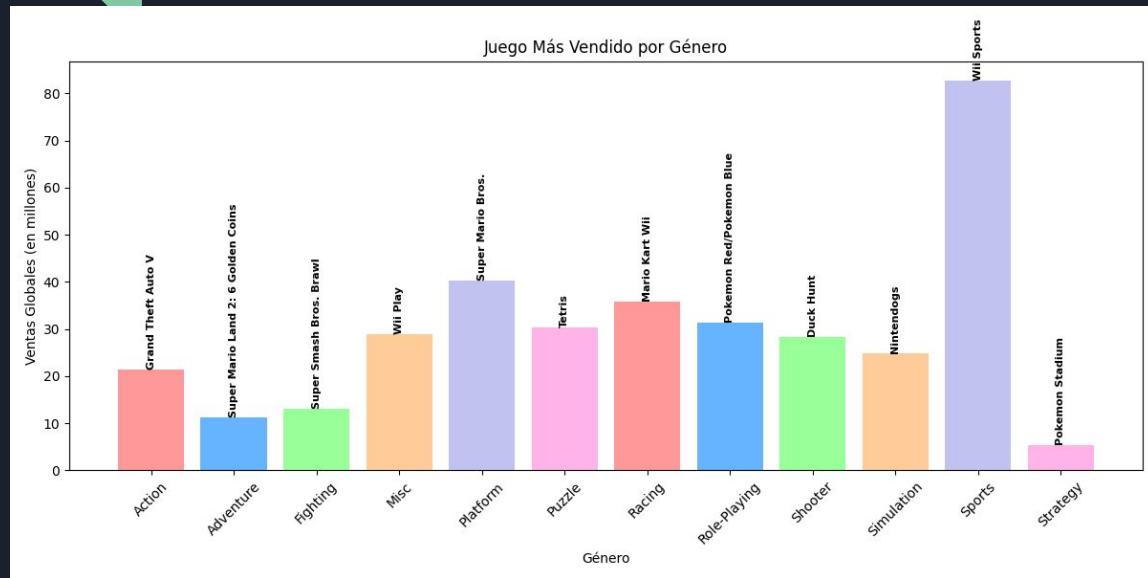
A partir del gráfico de ranking, se identificó que la empresa que obtuvo la mayor cantidad de ganancias a nivel mundial entre 2000 y 2010 fue Electronic Arts, seguida por Nintendo en segundo lugar y Activision en tercer lugar. Estas tres empresas fueron líderes en la industria de videojuegos durante esa década, con Electronic Arts ocupando la posición más alta en términos de ganancias acumuladas. Este análisis proporciona una visión clara sobre las empresas dominantes en la industria de los videojuegos durante ese período de tiempo específico.

### 3- ¿Cuál es el juego más vendido de cada una de las principales empresas?



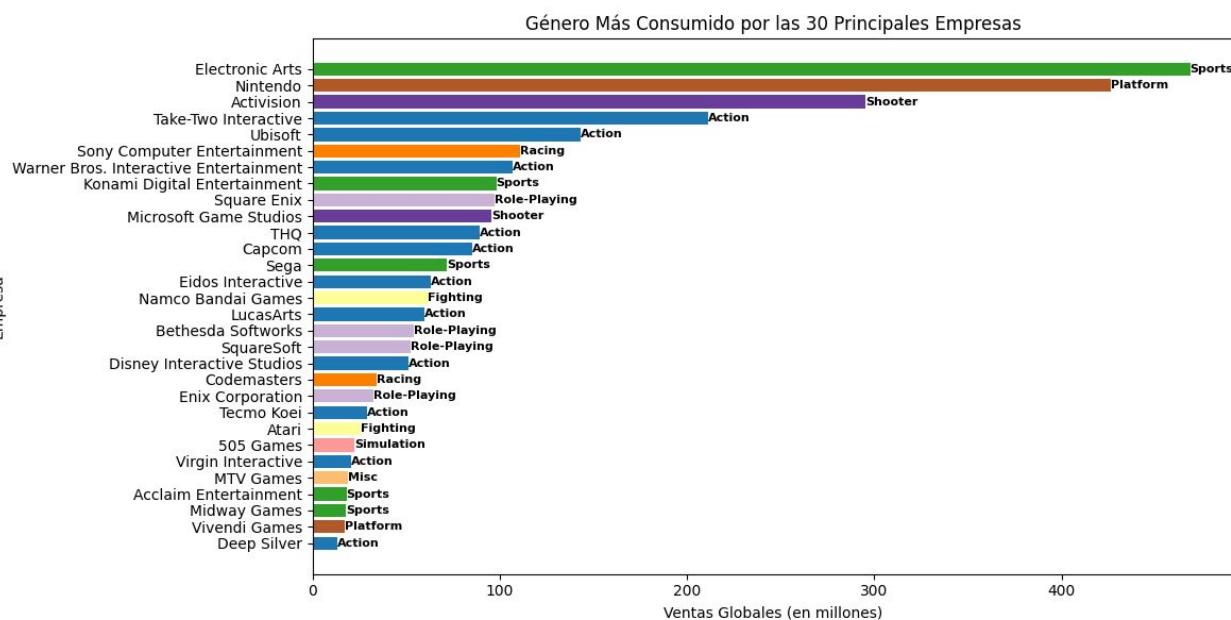
A partir del gráfico de ranking realizado se identificaron los juegos más vendidos de las principales empresas de videojuegos hasta 2016. Para Nintendo, el juego más vendido fue 'Wii Sports', destacándose como líder en ventas para la empresa. En el caso de Take-Two Interactive, 'Grand Theft Auto V' se posicionó como el título más vendido, subrayando el impacto duradero de esta franquicia en la industria. Para Activision, 'Call of Duty: Black Ops' se destacó como el juego más vendido, mostrando la popularidad continua de la serie 'Call of Duty'. Este análisis revela los títulos emblemáticos que han contribuido significativamente a las ventas de estas empresas, ofreciendo una visión clara de los juegos que han capturado la atención y el interés de los jugadores a lo largo de los años.

## 4- ¿Cuál es el juego más vendido para cada tipo de género?



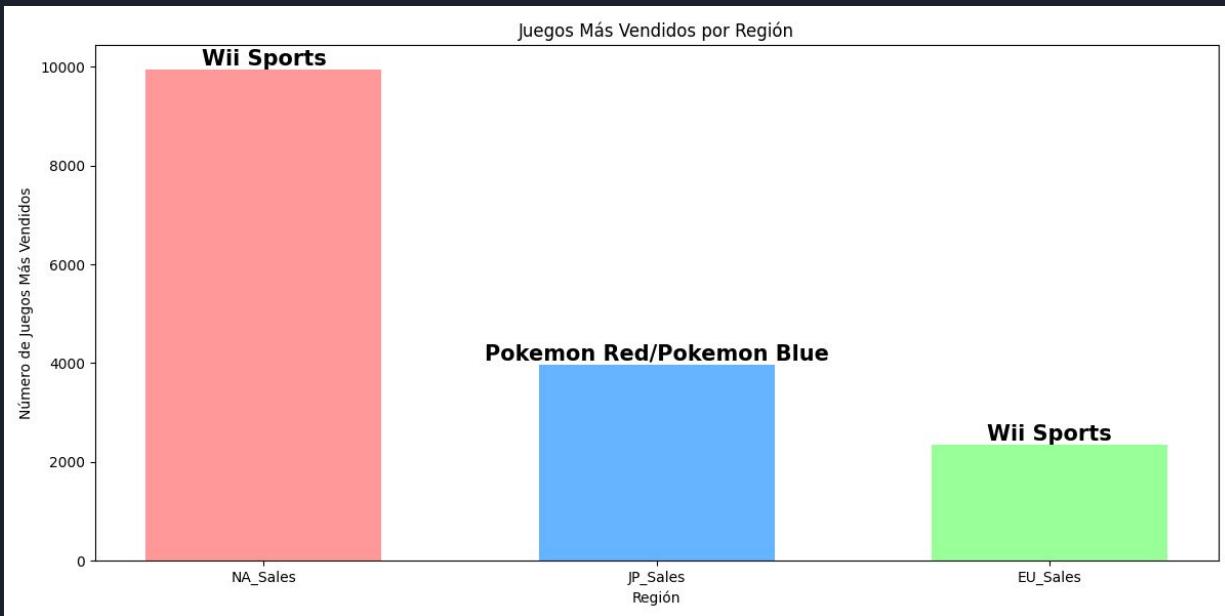
En términos de ventas globales, en el gráfico se visualiza que el juego 'Wii Sports' se considera el más consumido, perteneciendo al género Sports, de los demás videojuegos que se encuentran en el dataset. Las ganancias que se obtuvieron con este videojuego da una idea de las ventas que se pueden generar si una empresa produce un juego de género Sports. Le siguen a este videojuego: Super Mario Bros del género Platform y Mario Kart Wii de Racing

## 5- ¿Cuál fue el género de videojuego más consumido de cada empresa?



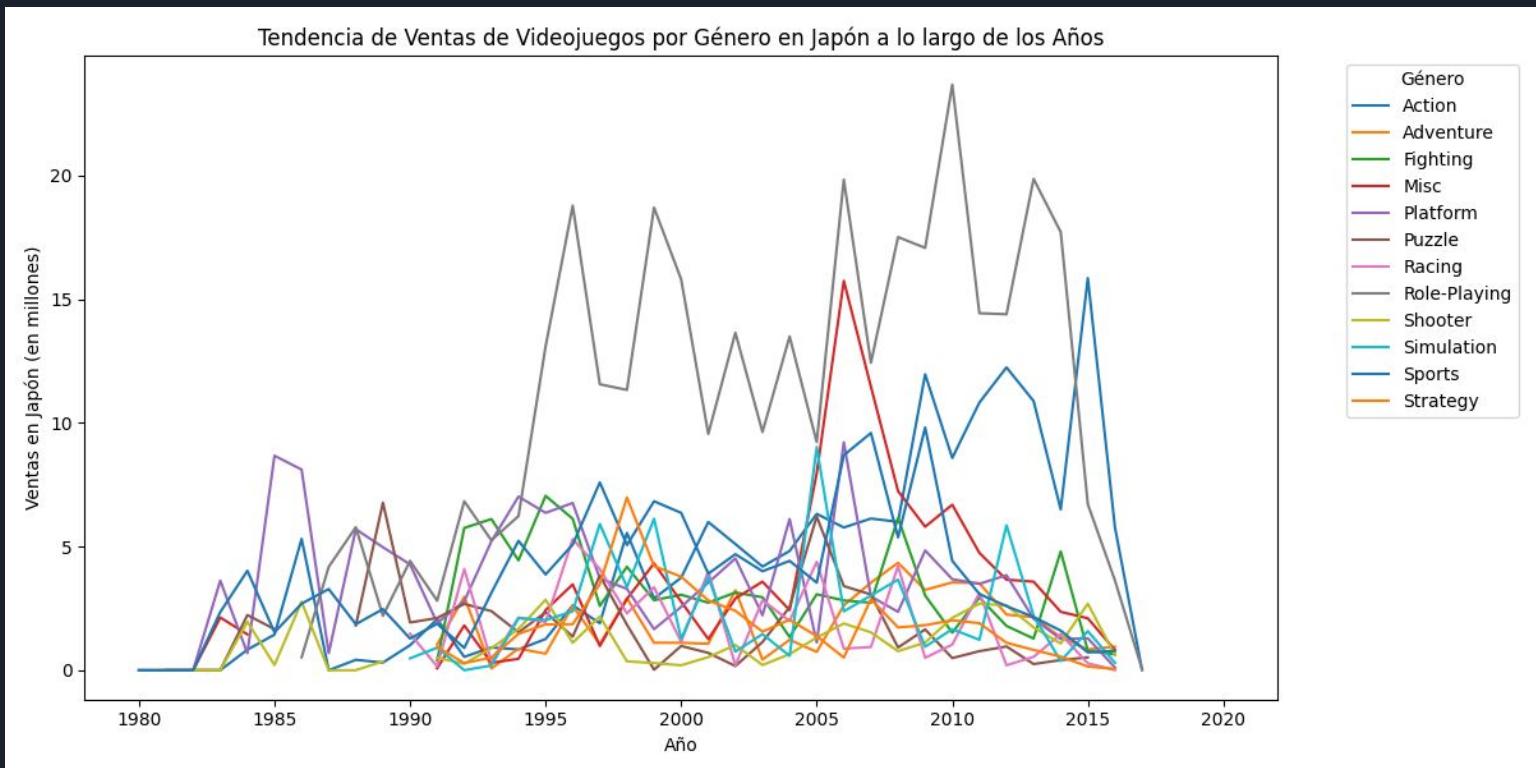
Se puede visualizar claramente que el género Sports sobresale del resto, siendo el más consumido en diferentes empresas. Esto deja en claro que los videojuegos de este género, junto a los de Platform y Shooter, son altamente consumidos y la producción de los mismos por parte de las grandes empresas podría generar altas ganancias en términos de millones

6- ¿Cuáles son los juegos más vendidos en cada región geográfica (por ejemplo, Norteamérica, Europa, Japón)?



La cantidad de ediciones que se vendieron para el juego Wii Sports está por encima de los demás videojuegos que se muestran en el gráfico. Dejando a la región de NorthAmerica como una de las regiones donde más se pueden llegar a vender videojuegos y obtener buenas ganancias

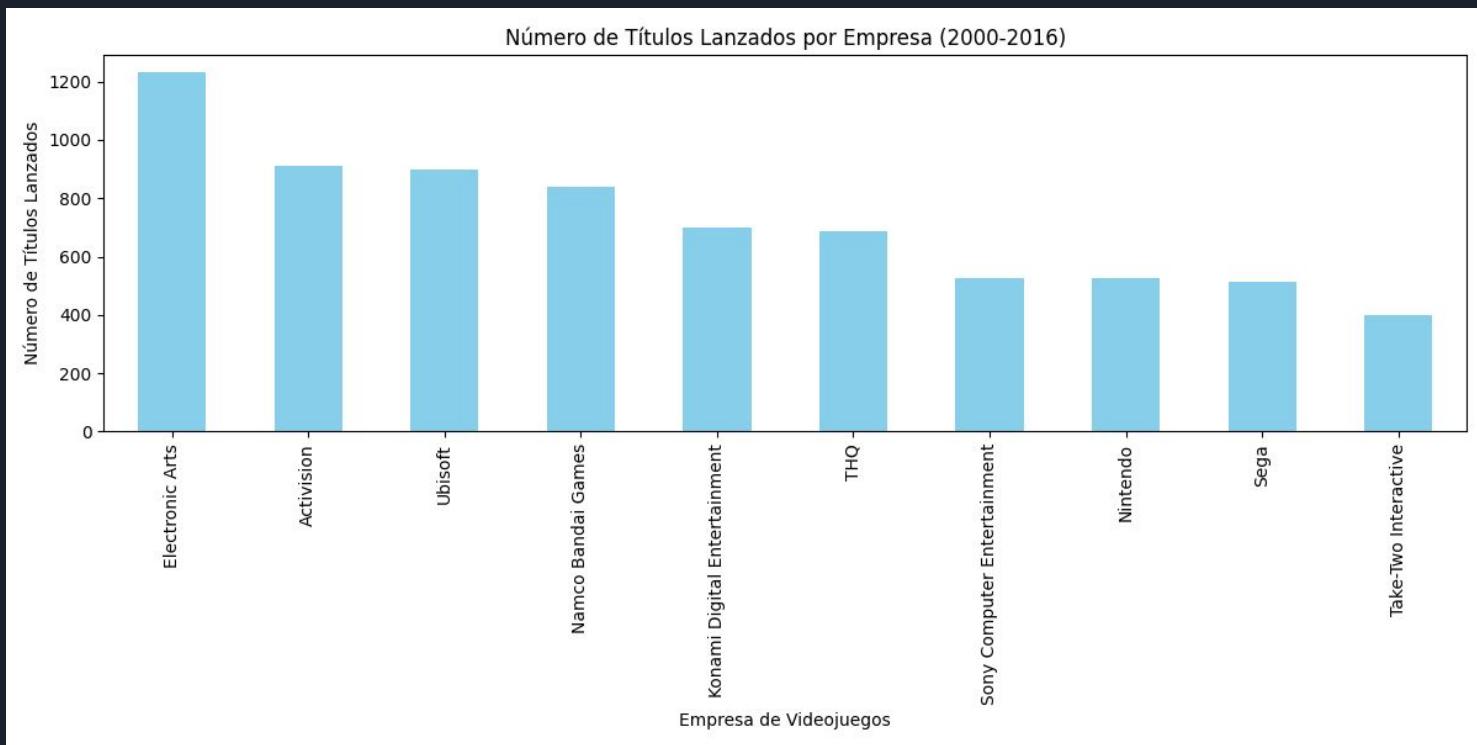
## 7- ¿Cuál es la tendencia de ventas de videojuegos por género en Japón a lo largo de los años?



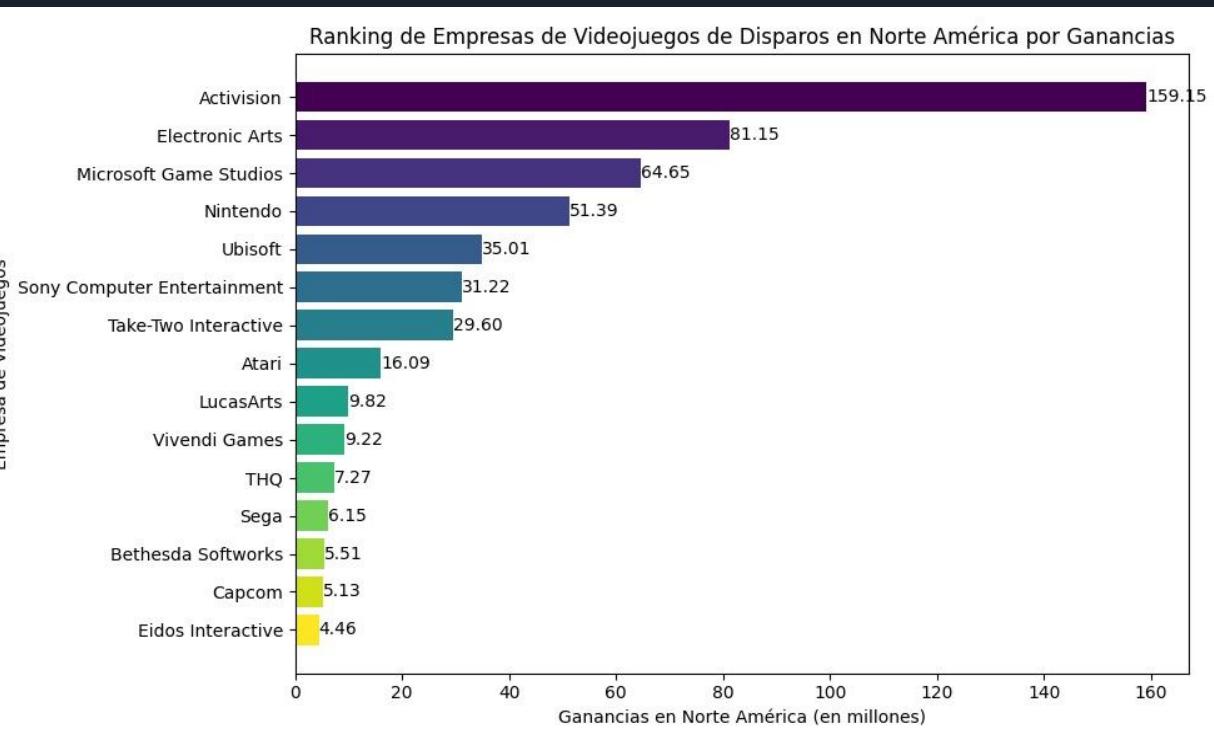
8- En qué año se registraron las ventas globales más altas en la historia de los videojuegos, y qué eventos o lanzamientos pueden haber contribuido a este logro?



## 9- ¿Cuál es la empresa de videojuegos con el mayor número de títulos lanzados desde 2000 a 2016?



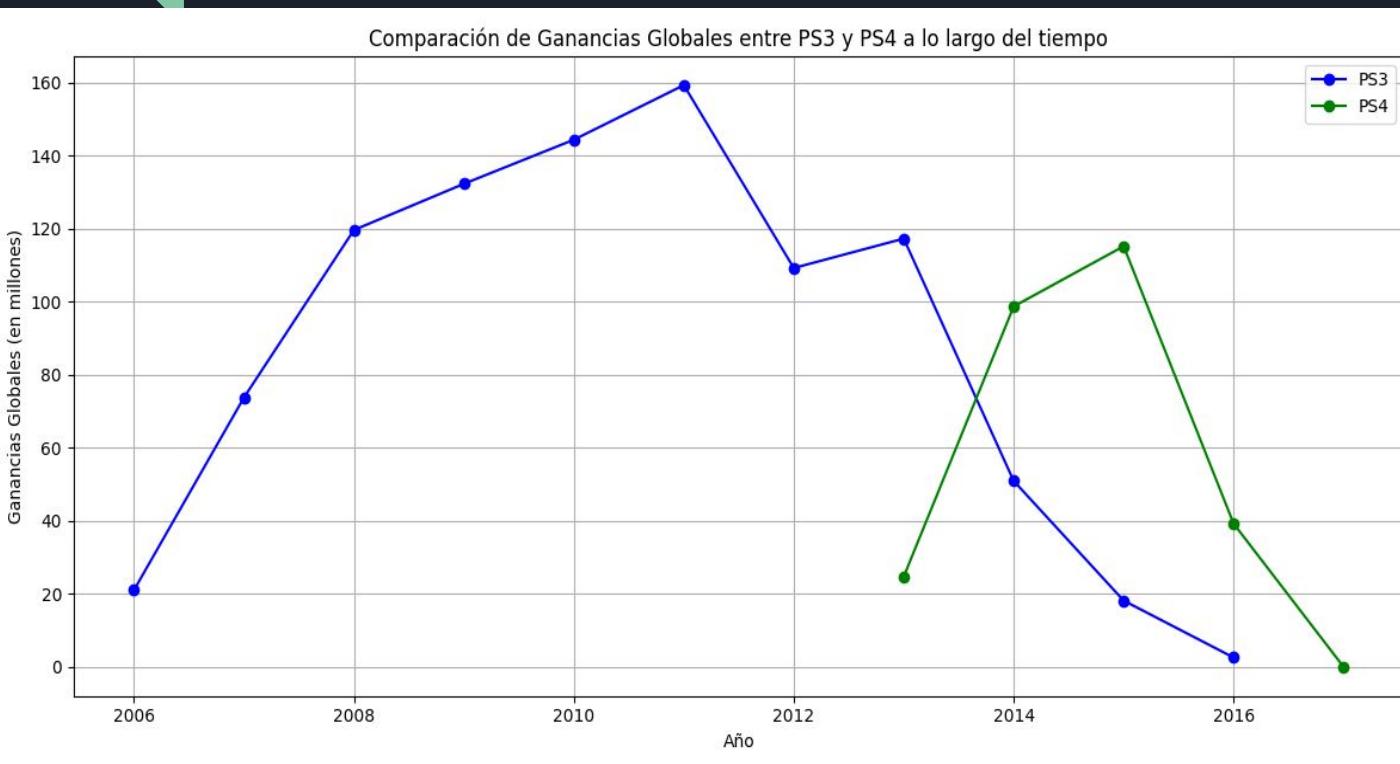
# 10- ¿Cuáles son las 15 principales empresas que obtuvieron mayor cantidad de ganancias en la región de Norte América para videojuegos tipo "Shooter"?



Para responder a esta pregunta, se analizó y se creó un gráfico de ranking que muestra las empresas de videojuegos de disparos ordenadas por ganancias en Norte América.

A partir del gráfico de ranking, se identificó que la empresa "Activision" fue la que obtuvo la mayor cantidad de ganancias en la región de Norte América para videojuegos tipo "Shooter".

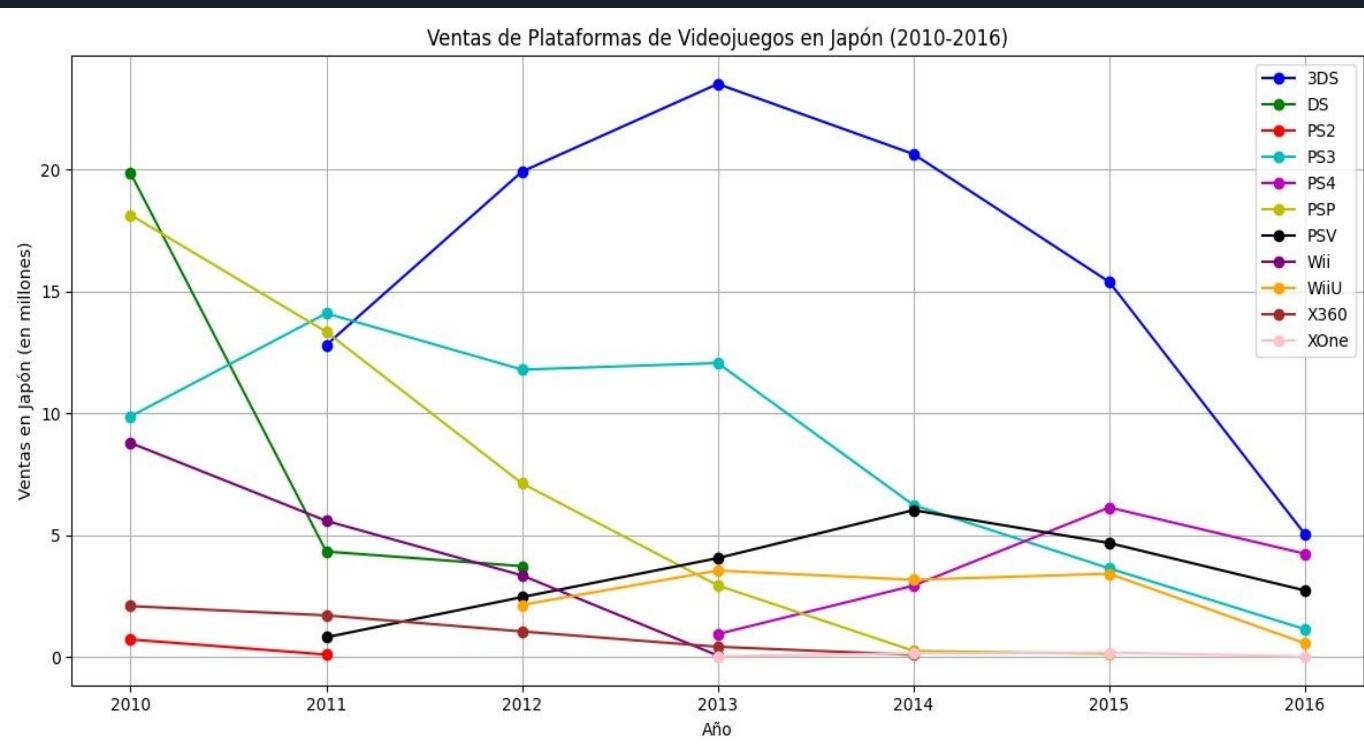
11- Luego de que salga a venta la PS4, en base a las ganancias globales, ¿Tuvo algún impacto negativo la plataforma PS3?



Se realizó un análisis de cambio a lo largo del tiempo (Gráfico: Change v Time) que comparó las ganancias globales de las plataformas PS3 y PS4 después del lanzamiento de la PS4.

A través del gráfico, se observó que las ganancias de la plataforma PS3 comenzaron a disminuir gradualmente después del lanzamiento de la PS4, lo que sugiere un posible impacto negativo en las ganancias globales de la PS3.

## 12- Tomando de referencia las ventas en Japón, ¿qué plataforma tuvo un impacto positivo en los últimos 6 años (2010-2016)?



Se creó un gráfico de cambio a lo largo del tiempo que muestra las ventas de diferentes plataformas en Japón a lo largo de los últimos 6 años (2010-2016).

A partir del análisis del gráfico, se identificó que “3DS” fue la que tuvo un mayor impacto positivo en los últimos 6 años, pero últimamente todas por igual terminaron en decadencia.



# Conclusiones

**Preferencias de Género y Empresas:** Las preferencias de género varían según la región y la empresa de videojuegos. Por ejemplo, en Estados Unidos, Nintendo tiene éxito en géneros como "Action" y "Sports". Esto sugiere que las empresas deben adaptar sus estrategias de desarrollo y marketing según la región y el género de juego.

**Líderes en Ganancias Globales:** Durante el período de 2000 a 2010, Electronic Arts, Nintendo y Activision fueron las empresas líderes en ganancias a nivel mundial. Esto indica que estas empresas tuvieron un impacto significativo en la industria durante esa década.

**Títulos Emblemáticos:** Los juegos más vendidos de cada empresa son títulos emblemáticos que han contribuido de manera significativa a las ventas. Ejemplos incluyen 'Wii Sports' para Nintendo y 'Grand Theft Auto V' para Take-Two Interactive.

**Impacto de las Plataformas:** El análisis de las ventas en diferentes plataformas muestra cómo el lanzamiento de nuevas consolas puede afectar las ganancias de las plataformas más antiguas, como la transición de PS3 a PS4.



**Tendencias de Ventas en Japón:** El género "Role-Playing" muestra una tendencia positiva en ventas en Japón a lo largo de los años. Esto podría reflejar las preferencias de los jugadores japoneses.

**Años de Ventas Globales Más Altas:** 2008 fue el año con las ventas globales más altas en la historia de los videojuegos. Los eventos o lanzamientos significativos pueden haber contribuido a este logro.

**Número de Títulos Lanzados:** Electronic Arts, Activision y Ubisoft lanzaron la mayor cantidad de títulos entre 2000 y 2016, lo que refleja su actividad constante en el mercado de videojuegos.

**Impacto de Plataforma en Japón:** La plataforma "3DS" tuvo un impacto positivo en las ventas en Japón entre 2010 y 2016, lo que sugiere que las nuevas plataformas pueden revitalizar el mercado.

Para concluir, el análisis de datos en el mundo de los videojuegos proporciona información valiosa para las empresas y desarrolladores al comprender las preferencias regionales, el impacto de las consolas, y la influencia de títulos icónicos. Estos datos pueden orientar estrategias de desarrollo, marketing y toma de decisiones en la industria de los videojuegos.