

## Pregunta 1

He elegido un dataset relacionado con videojuegos que contiene información muy relevante, especialmente sobre las ventas obtenidas a nivel global y por regiones individuales. Este conjunto de datos es óptimo para realizar un análisis de datos, ya que permite investigar, por ejemplo, las ventas en determinados años, identificar cuáles son los juegos más populares según su género y determinar en qué plataformas se registran más ventas.

A continuación, describo las columnas y la estructura de mi dataset, que consta de 11 columnas y aproximadamente 16,599 filas:

1. **Rank:** Número de clasificación del juego en la lista.
2. **Name:** Nombre del videojuego.
3. **Platform:** Consola o plataforma en la que se lanzó el juego (por ejemplo, PS3, GB, WiiU, etc.).
4. **Year:** Año en que se lanzó el juego.
5. **Genre:** Género del juego (por ejemplo, Sports, Platform, Action, etc.).
6. **Publisher:** Compañía que publicó el juego.
7. **NA\_Sales:** Ventas en millones de unidades en Norteamérica.
8. **EU\_Sales:** Ventas en millones de unidades en Europa.
9. **JP\_Sales:** Ventas en millones de unidades en Japón.
10. **Other\_Sales:** Ventas en millones de unidades en otras regiones.
11. **Global\_Sales:** Ventas globales del juego en millones de unidades.

En este dataset, la variable objetivo es Global\_Sales, ya que representa las ventas totales de cada videojuego en todo el mundo. Al analizar otras columnas, como el género del juego y la compañía que lo publicó, etc. podemos observar cómo estas características impactan en las ventas globales. Este tipo de análisis nos permite detectar tendencias y patrones en las ventas, lo cual es esencial para comprender el desempeño de los videojuegos en el mercado.