



Universidad del Rosario

Escuela de Ingeniería, Ciencia y Tecnología

## Segunda entrega

Prince Nicolas Bermudez Tabares

Sergio Andrés Guevara Ramírez

Maria Alejandra Rico León

Samuel Mauricio Scarpati Sarmiento

11 de Mayo de 2024

# 1. Correcciones Según la Primera Entrega

## 1.1. Pregunta Identificada

Los videojuegos son un componente de suma relevancia dentro del contexto del entretenimiento mundial. Es por eso que el objeto principal de estudio de este proyecto son los videojuegos, pero el estudio de los videojuegos como concepto, de por si es bastante amplio, ya que puede abordarse desde bastantes perspectivas dependiendo los aspectos que se quieran analizar.

El caso puntual de estudio que se abordará en este proyecto gira en torno al impacto financiero de la industria de los videojuegos, donde el contexto del proyecto el término 'impacto financiero' se reduce a una métrica de desempeño y calidad según el número de ventas que uno o muchos videojuegos pueden tener alrededor del mundo ó en regiones puntuales consideradas regiones de mayor impacto.

Se busca, que a traves de diversas herramientas del análisis de datos, las cuales serán mencionadas y exploradas en profundidad más adelante. Se identifique los actores fundamentales que dan lugar a la existencia, uso y clasificación de los videojuegos, y así alcanzar una conclusión de quienes tienen los mejores resultados, basandose como ya se menciono anteriormente, en los atributos de ventas que posee cada videojuego. Finalmente, para conseguir esto, es necesario un conjunto de datos sobre el cual trabajar por lo cual se toma el conjunto de datos ***Video Games Sales***, a través de la plataforma

en pro de la Ciencia de Datos llamada **Kaggle**, donde por medio de un proceso de *web scraping* sobre la página <https://www.vgchartz.com/#>, se nos presenta información como el año de lanzamiento del videojuego, su compañía desarrolladora, el título del videojuego, plataformas, generos y al menos 5 componentes de información de ventas distintos. Que serán más que suficientes para el trabajo pertinente.

Acceso al conjunto de datos en cuestión : <https://www.kaggle.com/datasets/gregorut/videogamesales>

## 1.2. Entidades

- **Videojuego** : El elemento central de los datos, quien existe debido a una desarrolladora, hecho para una o más consolas (ó plataformas).
- **Desarrolladora** : Compañía quien desarrolla el proceso técnico de la creación del videojuego y la publicación de este mismo. .
- **Consola** : Medio por el cual se hace uso de uno o más videojuegos.
- **Venta** : Acción de trasferencia del producto (videojuego), al cliente final (jugador).
- **Genero** : Clasificación otorgada a un videojuego, según el contenido del videojuego mismo.

### 1.3. Reglas de Negocio

- Cada videojuego debe tener únicamente una compañía desarrolladora.
- Solamente puede haber un genero por cada videojuego.
- No puede haber mas de un año de lanzamiento por cada videojuego.
- Un videojuego debe de tener una o más plataformas asociadas.
- Cada desarrolladora debe tener a su nombre como mínimo un videojuego.
- Una desarrolladora  $k$ , no puede figurar como desarrolladora de un videojuego  $V$ , si este ya tiene una desarrolladora diferente a  $K$ , asignada.
- Todas las desarrolladoras tienen nombre.
- Cualquier desarrolladora puede realizar videojuegos bajo cualquier genero.
- Toda plataforma o consola, debe tener una sola compañía creadora responsable.
- Toda plataforma o consola, necesita una fecha de lanzamiento, a excepción de la plataforma de PC, ya que esta no es de por si un producto de una única empresa.
- Cualquier compañía dueña de cualquier plataforma o consola, puede tener a su nombre una o más plataformas distintas.

- Todas las plataformas o consolas tienen nombre.
- Las ventas totales de un videojuego son fragmentadas en ventas NA(Norte América), ventas JP(Japón), ventas EU(Unión Europea) y otras ventas.
- Las ventas de cada videojuego estan dadas en millones.
- Cada venta representa una copia del titulo(videojuego) vendida.
- Por cada videojuego debe de haber al menos una copia vendida.
- Puede haber cuanta cantidad de generos sean necesarios.
- No puede haber más de un genero asociado a un videojuego.
- Si se contempla un conjunto  $X$  de géneros, en el caso de observar todos los videojuegos emparejados con su respectivo genero, debe haber al menos un videojuego asociado para cada  $x$  en  $X$ .
- No pueden haber generos sin nombre.
- Cualquier Titulo(videojuego) puede tener una o más versiones, dependiendo de sus plataformas asociadas.
- Los años de lanzamiento entre versiones de un mismo videojuego pueden o no ser todos iguales.

## 1.4. Modelo Entidad - Relación

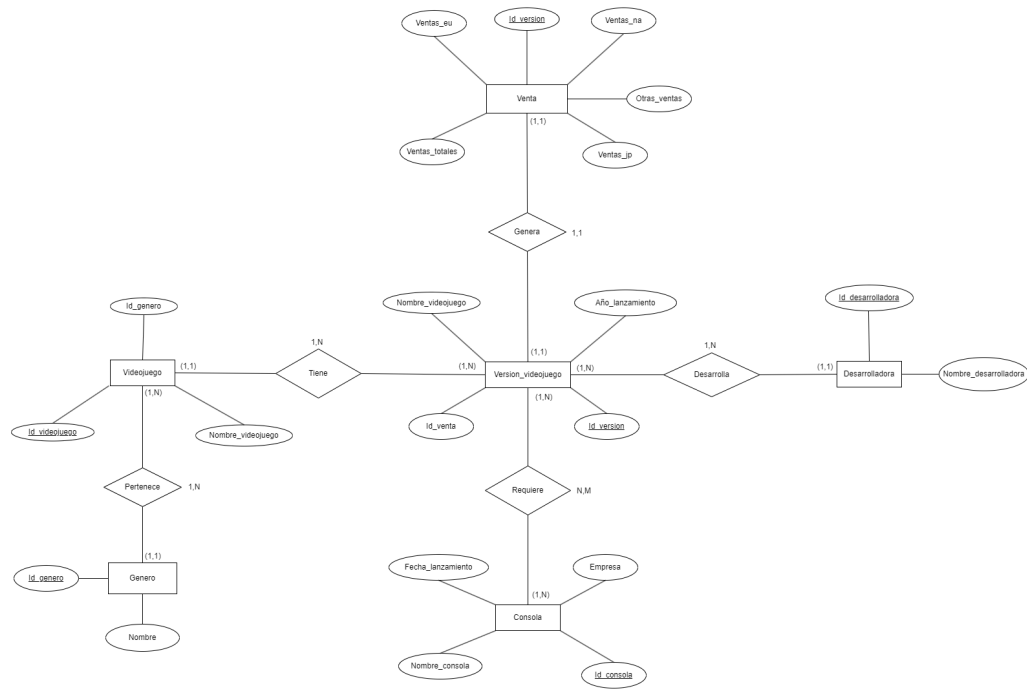
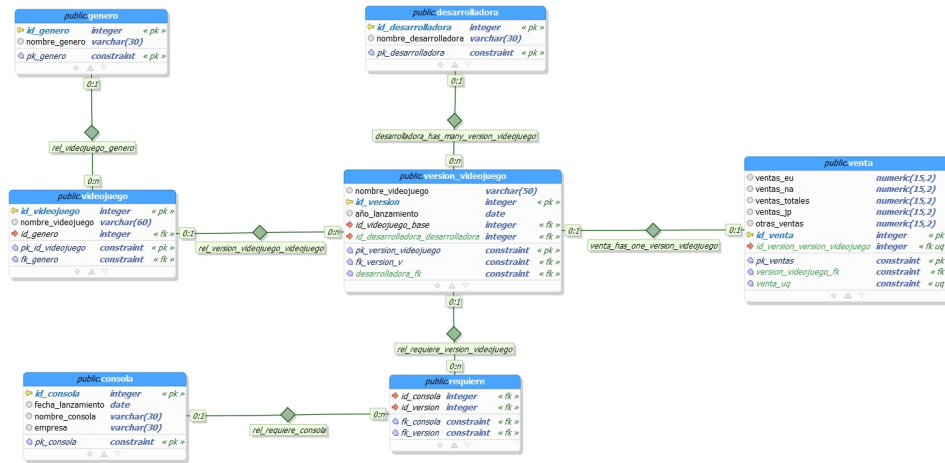


Figura 1: Modelo Entidad - Relación

## 1.5. Diagrama Relacional



Activar Window

Figura 2: Diagrama Relacional

## 2. Escenarios de Análisis

- a) Por cada desarrolladora , determinar cual es el género , para el que más han desarrollado videojuegos, y para el que menos han desarrollado videojuegos. Para así tomar el contraste que existe sobre la cantidad de unidades vendidas entre el genero con mas desarrollo y el género con menos desarrollo.

Este caso analiza la relación entre el género de los videojuegos desarrollados por cada compañía y las unidades vendidas.

- b) Filtrar para toda la base de datos el género con mas videojuegos desarrollados, y realizar un promedio de copias vendidas desde que se lanzo el primer juego del género, hasta la fecha actual.

Este caso busca analizar el rendimiento de ventas de un género específico a lo largo del tiempo. Es útil para comprender la evolución y el éxito relativo de un género en particular.

- c) Determinar cual ha sido el videojuego con el mayor número de ventas totales y la plataforma para la cuál se de este escenario.

Este caso se enfoca en identificar el videojuego más exitoso en términos de ventas totales y la plataforma asociada. Es útil para comprender qué juegos y plataformas han sido más populares entre los consumidores.

- d) Encontrar los 3 videojuegos con más versiones existentes para distintas plataformas y los 3 videojuegos con menos versiones existentes. Para



así realizar una comparación entre, la cantidad de ventas totales y la cantidad de versiones.

Este caso analiza la relación entre la cantidad de versiones de un videojuego y sus ventas totales. Es útil para comprender cómo la disponibilidad en múltiples plataformas afecta el rendimiento de un juego.

e) Consolidar 5 tablas de clasificación para las compañías desarrolladoras:

- Clasificación de desarrolladoras según ventas totales en Japón.
- Clasificación de desarrolladoras según ventas totales en Norte América.
- Clasificación de desarrolladoras según ventas totales en la Unión Europea.
- Clasificación de desarrolladoras según otras ventas.
- Clasificación de desarrolladoras según la suma total de ventas en Japón, ventas en Norte América, ventas en la Unión Europea y otras ventas.

Para así finalmente realizar una clasificación final donde veamos de la desarrolladora con el mayor desempeño promedio de ventas a la desarrolladora con el menor desempeño promedio de ventas.

f) Realizar un caso de estudio sobre si existe una relación entre el nombre de la desarrolladora y la cantidad de copias que venden sus juegos, es decir, analizar si la reputación de la desarrolladora la precede.

### **3. Creación de base de datos, carga de datos y desarrollo de módulos en Python**

Los procesos pertinentes para llevar cabo a estas acciones, se pueden encontrar en el repositorio de GitHub del proyecto, a través del siguiente link  
[https://github.com/Serkx423k/Proyecto\\_Ingenieria\\_De\\_Datos.git](https://github.com/Serkx423k/Proyecto_Ingenieria_De_Datos.git)