



Universidad del Rosario

Escuela de Ingeniería, Ciencia y Tecnología

Tercera Entrega

Prince Nicolás Bermúdez Tabares

Sergio Andrés Guevara Ramírez

Alejandra Rico León

Samuel Mauricio Scarpati Sarmiento

25 de mayo de 2024

Modelo Entidad - Relación

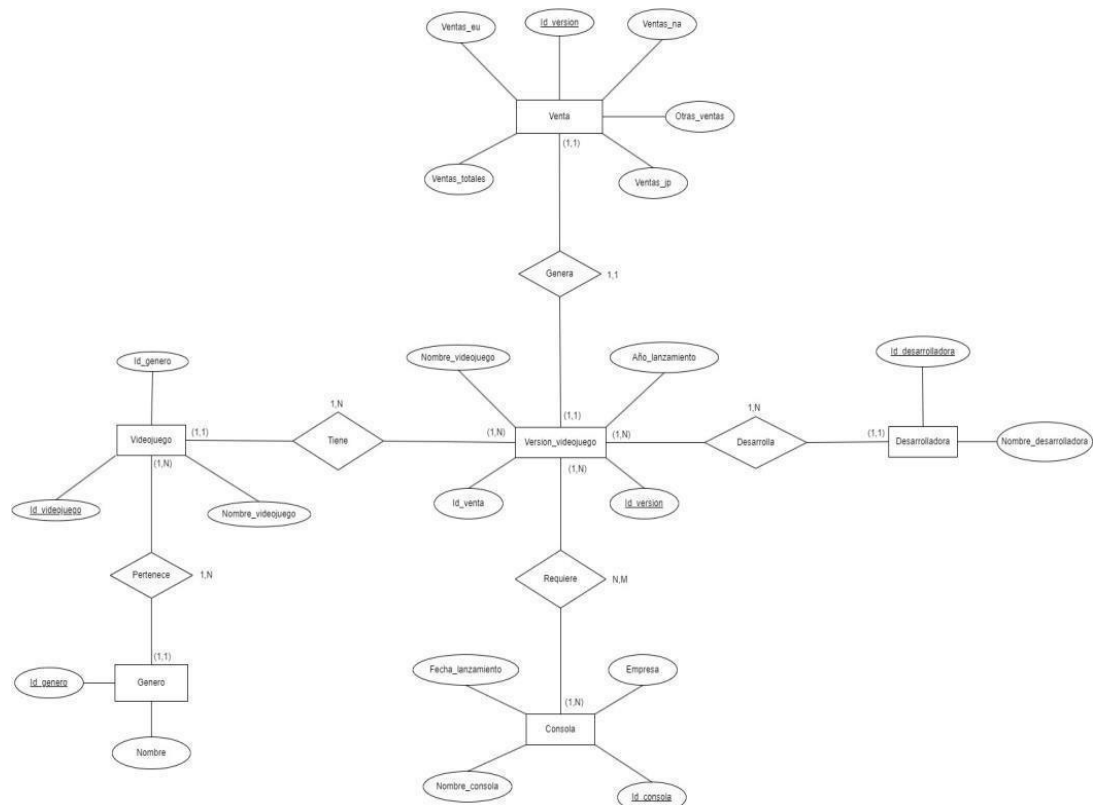


Figura 1: Modelo Entidad - Relación

Diagrama Relacional



Figura 2: Diagrama Relacional

Activar Window

Escenarios de Análisis

- (a) Por cada desarrolladora, determinar cuál 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 género con más 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) Encontrar la cantidad de juegos desarrollados por género, y realizar un promedio de copias vendidas por región de cada género. Este caso nos permite entender como el género de un juego afecta sus ventas en diferentes partes del mundo.
- (c) Hacer una gráfica sobre las ventas anuales del género con mayor número de juegos desarrollados. 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.
- (d) Determinar cuál ha sido el videojuego con más ventas totales y la plataforma para la que se da 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.

Discusión

En esta discusión, nuestro objetivo es identificar y evaluar las características de un análisis desarrollado que relaciona los géneros de videojuegos desarrollados por cada compañía con las unidades vendidas. Específicamente, analizaremos:

Prince Bermudez

- (a) Comparación de géneros: Muestra claramente la proporción de diferentes géneros desarrollados, pero no es ideal para comparar más de una desarrolladora a la vez o para comparar subcategorías dentro de los géneros.
- (b) Análisis de juegos por género y ventas promedio: Proporciona una visualización clara de la cantidad de juegos desarrollados por género. Pero puede ser confuso al momento de poner las regiones.
- (c) Ventas anuales de genero con mayor número de juegos desarrollados: El gráfico de línea es excelente para mostrar tendencias a lo largo del tiempo, permitiendo identificar picos y valles en las ventas anuales. Sin embargo, este tipo de datos no es ideal para comparar múltiples géneros.

- (d) Video juego con mayor número de ventas y su plataforma: Nos permite visualizar cuál fue el juego con mayores ventas, pero podría ser más claro si esto se visualizara en un top con estos juegos y sus plataformas.

Alejandra Rico

- a) Mediante este diagrama, podemos analizar las ventas de cada desarrolladora en base a cada género, es decir, con este se nos es posible determinar el género para el que más y para el menos se generaron ventas por cada compañía, lo que es útil a la hora de analizar las ventas de cada compañía y en el desarrollo de qué género tuvieron mayor auge entre las masas. Sin embargo, este diagrama deja por fuera las ventas hechas de los géneros diferentes al más y el menos vendido, lo que hace difícil generar un marco generalizado de las ventas de cada compañía.
- b) En el diagrama podemos visualizar el promedio de la cantidad de ventas por género en diferentes regiones del mundo. Este diagrama facilita la tarea de encontrar el género con mayor acogida entre la población de ciertas regiones y del mundo, además de proporcionar información acerca de la cantidad de juegos desarrollados por género, para así saber los géneros con mayor y menor desarrollo. La desventaja de este diagrama radica en que son dos diagramas diferentes que, pese a poder relacionarse, proporcionan diferente información y puede llegar a ser confuso al analizarlos.
- c) Este gráfico muestra las ventas globales del género Acción con el paso de los años. Este gráfico nos permite analizar la evolución que ha tenido el género Acción en el mercado global a través de los años, y así entender su éxito y auge entre la población. Sin embargo, este diagrama es insuficiente, pues solo nos proporciona la información de un género en específico, Acción, y deja totalmente de lado los otros géneros, evitando así el poder hacer un análisis anual de las ventas globales de cada género.
- d) Por último, la información proporcionada detalla el videojuego con mayor número de ventas totales, el cuál es Wii Sports, y la plataforma de dicho videojuego, la cual es Wii. Estos datos son importantes pues muestran el juego de mayor auge de los últimos años y su plataforma, lo cual es esencial al crear un panorama de las preferencias de los usuarios en las últimas décadas. Por otro lado, no nos brinda una información completa, pues se limita a darnos la información de solo el juego más vendido, dejando de lado todos los demás juegos desarrollados en los últimos años y su número de ventas totales.

Samuel Scarpati

- a) *Comparación de Géneros por Desarrolladora:* Para abordar este análisis, utilizamos una visualización de gráfico de pastel. En este gráfico, cada pastel representa a una desarrolladora, mostrando el género que más desarrollaron y el que menos, todo a través de un porcentaje y de la cantidad total de juegos.
- Ventajas: Este grafico es excelente en cuanto a la comparación directa, es sencillo identificar en un vistazo cual es la información, no solo en cuanto a lo

que desarrolló una sola compañía, sino también es rápido encontrar los datos de las demás compañías en la base de datos.

- Desventajas: Cuando la cantidad de juegos desarrollados es muy pequeña comparada con el dato mayoritario, puede ser difícil visualizar segmentos pequeños, lo cual puede llevar a una mala lectura o una interpretación ambigua. De igual modo, este gráfico no muestra de manera específica la causalidad de por qué un género puede ser más desarrollado que otros.
- b) *Análisis de juegos por género y ventas promedio:* Para este análisis, usamos un gráfico de barras que nos muestran ambas partes de la información consultada. Por un lado, se genera un gráfico con la información de juegos desarrollados por género; y por el otro, hay un gráfico con el promedio de copias vendidas por región.
- Ventajas: De nuevo, la comparación entre datos y géneros es muy fácil ya que todos los datos se muestran juntos y de manera clara, de igual manera, al ser fácil de leer se sabe cuáles son los mercados clave en la industria de los videojuegos, incluso, gracias a ambas consultas, saber cómo varían los mercados en diferentes partes del mundo.
 - Desventajas: Un problema común al promediar datos es que pueden ser influenciados por un pequeño número de juegos con ventas extremadamente bajas o altas, distorsionando los datos de éxito en una región específica.
- c) *Ventas anuales de género con mayor número de juegos desarrollados:* Nuestro enfoque es analizar el rendimiento de ventas anuales del género más destacado. Para observar esta tendencia usamos una gráfica de líneas, la cual es ideal para observar datos que cambian con el tiempo, en este caso podemos ver picos y crecimiento durante varias décadas para lograr un análisis completo.
- Ventajas: Como ya fue mencionado, una ventaja de una gráfica lineal es ver la transformación de los datos a lo largo del tiempo, facilitando el entendimiento de cómo ha cambiado la industria a través de los años.
 - Desventajas: Pese a que permite ver cómo han cambiado las tendencias, no muestra una clara razón para ello. En muchos casos el cambio en demanda puede ser por nuevas tecnologías, innovaciones o simplemente el público comienza a buscar cosas diferentes, sin embargo, no es posible concluir eso solo viendo las ventas.
- d) *Video juego con mayor número de ventas y su plataforma:* En este caso solo hacemos una exhibición de los datos tomados de las ventas de juegos en todas las consolas, de esta manera, vemos cual consola ha sido la más exitosa, no en sus ventas, sino en el uso de su plataforma.
- Ventajas: Esta visualización es importante para entender que las ventas no siempre se tratan de los juegos, sino de la variedad y exclusividad que puede ofrecer una plataforma a sus usuarios, de este modo se proporciona una visión completa de los escenarios de éxito.
 - Desventajas: Enfocarse en un solo juego con el mayor número de ventas limita la comprensión del panorama general de la industria.

- a) *Comparación de Géneros por Desarrolladora*: Este escenario de análisis se centra en contrastar los géneros de videojuegos más desarrollados y menos desarrollados para cada desarrolladora presente en la base de datos.

Esta visualización se hace realidad por medio de un gráfico circular que indica la cantidad de juegos que se han hecho, y el porcentaje que esta cantidad representa para la cantidad total de juegos desarrollados.

- Ventajas: Al solo haber 2 aspectos a comparar (genero más y menos desarrollado) un diagrama circular es una manera inmediata y bastante fácil de interpretar, además, la relación visual con los porcentajes representa una gran ayuda a la comprensión del gráfico.

-Desventajas: Una desventaja notoria, es que para un gráfico circular que cuenta con las etiquetas de información sobre los segmentos del gráfico, si el porcentaje de este es muy pequeño, la información puede no visualizarse.

- b) *Análisis de juegos por género y ventas promedio*: Esta sección de análisis, se segmenta en dos secciones principales.

La primera: Se puede visualizar un diagrama de barras que indica un panorama general sobre cuál es el género de videojuegos para el que más videojuegos se han hecho, así también como para los generos que menos videojuegos se han hecho.

La segunda: Siguiendo el tópico de los géneros. En esta segunda parte se presenta un diagrama de barras donde según el género que se escoja, se visualiza la comparación de los promedios de ventas norte américa, unión europea, otras ventas y la suma de todos los parámetros de ventas denominada “ventas globales promedio”.

- Ventajas: Es útil determinar cuál ha sido el género con más desarrollo mediante un diagrama de barras para futuros análisis y también para ver las relaciones entre los promedios de ventas que hay para el género consultado.

- c) *Ventas anuales de género con mayor número de juegos desarrollados*:

Esta sección contiene una gráfica de línea de tiempo, para el género con mayor número de videojuegos desarrollados, en donde se relaciona el número de ventas del videojuego, con un año en específico.

- Ventajas: Este modo de visualización, es bastante útil para para comprender el rendimiento del género para el que más juegos hacen a lo largo del tiempo.
- Desventajas: Sería aún mejor, si en vez de visualizar el rendimiento a lo largo del tiempo únicamente para el género más desarrollado, pudiera apreciarse también para otros géneros.

- d) *Video juego con mayor número de ventas y su plataforma*: El ultimo escenario de análisis presente en la aplicación web muestra el nombre del videojuego más vendido, el número de copias vendidas, y para que plataforma se vendieron esta cantidad de copias.

- Ventajas: Es bastante sencillo, pero a la vez muy importante, ya que brinda un veredicto sobre la versión de videojuego más definitiva, y esto más que solo

hablar del juego en sí, también brinda bastante información sobre el posicionamiento de las consolas o plataformas.

- Desventajas: Al ser tan concisa la sección brinda poca información, lo que también hace prescindir del uso de medios de representación gráfica, lo que difiere con la estructura mayoritaria de la aplicación.

Conclusiones

Prince Bermudez

Para la creación de la interfaz de usuario, optamos por Dash, lo que nos permitió desarrollar dashboards interactivos y visualizaciones dinámicas. La integración con Plotly facilitó la generación de gráficos que mostraban tendencias y comparaciones de ventas por región y género. La optimización del rendimiento de la aplicación fue un desafío significativo que abordamos mediante técnicas de optimización de consultas y mejoras en la eficiencia de la aplicación. Utilizamos herramientas como pgAdmin y pgModeler para diseñar nuestra base de datos, comenzando desde la primera forma normal (1NF) y avanzando hacia un modelo de datos completamente normalizado. Este enfoque estructurado permitió reducir la redundancia y mejorar la integridad de los datos, facilitando consultas eficientes y una gestión de datos optimizada. La primera fase del proyecto implicó la identificación y selección de fuentes de datos confiables sobre las ventas de videojuegos en diversas regiones (EU, JP, NA y globales) y géneros. La calidad y consistencia de los datos fueron factores críticos que aseguraron una base sólida para el análisis subsecuente. A pesar de los desafíos iniciales de limpieza y validación de datos, logramos consolidar una base de datos precisa y detallada.

Alejandra Rico

Con diferentes herramientas como HTML y CSS, se pudo desarrollar la página web donde está la información de nuestra base de datos, organizada y concisa para el mejor entendimiento del usuario. Para la creación de los gráficos mostrados en la página web, se usaron librerías de Python como Dash, Plotly y Pandas, cuyo entendimiento fue un desafío para nosotros, puesto que no teníamos ningún tipo de referencia o base para guiarnos y no fue un tema muy profundizado durante las clases, por lo que fue necesario comprender su funcionamiento de manera autónoma. Sin embargo, estas librerías nos facilitaron de sobremanera la construcción de todos los diagramas y, después de entender mucho mejor su funcionamiento, fue mucho más sencillo su manejo y manipulación. Por otro lado, para la creación del MER y del diagrama relacional, usamos el pgAdmin y el pgModeler, herramientas donde producir dichos diagramas era bastante sencillo, sin embargo, el problema para nosotros fue la organización de nuestra base de datos y, la creación de las tablas que íbamos a usar y el establecimiento de las relaciones entre cada una de ellas. Al superar estas dificultades y corregir los errores de la primera entrega, se pudo crear una página web funcional y útil para entender el auge de diferentes videojuegos según su género, desarrolladora, ventas en ciertas regiones, ventas globales, entre otros aspectos.

Samuel Scarpati

Para llevar a cabo nuestro proyecto de análisis, hemos utilizado SQL para consultar y gestionar los datos almacenados en una base de datos de PostgreSQL, que obtuvimos en Kaggle y exportamos desde archivos .csv. PgAdmin fue una herramienta esencial en el manejo y supervisión de la base de datos, allí realizamos todas nuestras consultas, mientras que pgModeler facilitó el modelado y diseño de la planeación y esquematización de la base de datos. La integración de estas herramientas con Python y bibliotecas como pandas nos permitió manipular los datos de manera eficiente, por último, la creación de material visual e interactivo con Dash y Plotly le proporciona al usuario una experiencia dinámica y cómoda. En nuestro trabajo, tuvimos obstáculos con la simplificación de la base de datos, originalmente, la información estaba toda conjunta, lo que es problemática al evitar redundancias y tener datos consistentes; pero logramos esparcir los datos para que el sistema tuviera buena escalabilidad y buen rendimiento. Estos análisis nos dan una visión integral del mercado de videojuegos: poder identificar tendencias en géneros y regiones, el seguimiento de la evolución de las ventas a lo largo del tiempo y la determinación de los títulos y consolas más exitosas es esencial para tomar decisiones informadas, y aún más importante, estratégicas.

Sergio Guevara

El proyecto en cuestión deja una basta lista de aprendizaje, ya que es un proceso completo en sí, empezando desde buscar un conjunto de datos apto sobre el cual trabajar, muestra que no es tan sencillo. La importancia de “preparar el terreno” es un aspecto muy importante, donde “preparar el terreno” significa que no se puede no se puede pasar directamente a crear una base de datos e insertar datos a diestra y siniestra, es muy importante preparar los datos bajo una forma estandarizada, que evite problemas como redundancia, etc. El diseño de la base de datos también es de suma importancia, ya que es el paso de preparación teórico más importante, que permitirá preparar la base de datos de la mejor manera para trabajar después sobre los datos.

También hay que mencionar lo importante que pueden ser la herramienta web junto con lenguajes de programación como Python, lenguajes de manipulación de bases de datos como SQL. Todo esto junto crea un equipo interesante, desde la extracción de los datos necesarios, hasta como se muestran estos al usuario.

Finalmente, la enseñanza de este proyecto es que no es lo más importante que datos usemos, lo realmente importante es que usos les damos a través de las herramientas disponibles.

Repositorio

Todos los procesos, archivos y documentación acerca del proyecto se encuentran en el repositorio en GitHub del proyecto, al cual se accede con el siguiente enlace:

https://github.com/SerKx423k/Proyecto_Ingenieria_De_Datos