

Proyecto Integrado (EA1). Formulación de una necesidad de ingeniería de datos

Mateo Lara Aristizábal

Simón Lara Aristizábal

IUDigital

Proyecto Integrado V - Linea de Énfasis PREICA2502B020074

Andrés Felipe Callejas

8 de noviembre de 2025

Contenido

Resumen.....	3
Objetivos del Proyecto	3
Objetivo General	3
Objetivos Específicos	3
Metodología	4
Definición y Preparación	4
Desarrollo y Análisis de Datos	4
Conclusiones y Documentación	4
Resultados	4

Resumen

El presente estudio tiene como objetivo principal realizar un análisis exhaustivo de las ventas históricas de videojuegos a nivel mundial, desglosando el rendimiento comercial a través de tres dimensiones clave: plataformas, géneros y regiones geográficas. Se busca identificar las tendencias predominantes en el mercado, cuantificando la contribución de cada plataforma de *hardware* (e.g., PS4, Xbox, PC) y de cada categoría de *software* (e.g., Acción, Deportes, Rol) al total de ventas globales. La investigación se centra en regiones específicas como Norteamérica (NA), Europa (EU), Japón (JP) y el resto del mundo (Other) para comprender las dinámicas culturales y económicas que impulsan el consumo. Este análisis inicial es fundamental para establecer un marco de referencia cuantitativo que permita a los desarrolladores, editores y analistas de mercado tomar decisiones estratégicas informadas. Los hallazgos proporcionarán una visión clara sobre dónde y cómo se ha generado el mayor valor económico en la industria del videojuego a lo largo de los años. Este trabajo se basa en datos históricos de ventas y utiliza técnicas de minería de datos y programación Python con SQLite para su procesamiento y visualización.

Objetivos del Proyecto

Objetivo General

Analizar las ventas de los videojuegos por plataformas, géneros y regiones, con el fin de identificar las tendencias históricas de consumo y las dinámicas de mercado predominantes en la industria global.

Objetivos Específicos

Determinar el rendimiento de ventas de cada plataforma a lo largo del tiempo, identificando las consolas o sistemas de mayor impacto global y regional.

Cuantificar la distribución de las ventas de videojuegos por género en las principales regiones (NA, EU, JP), estableciendo qué categorías de *software* generan la mayor cuota de mercado en cada área.

Comparar el comportamiento de las ventas entre las regiones (Norteamérica, Europa, Japón y Otras) para destacar las diferencias culturales y preferencias de consumo específicas.

Establecer las herramientas y procedimientos de extracción, transformación y carga necesarios para migrar el dataset de ventas a una base de datos relacional (**SQLite**) utilizando Python.

Metodología

Definición y Preparación

Desarrollo y Análisis de Datos

Conclusiones y Documentación

Resultados