



INFORME DEL ANÁLISIS DE BASE DE DATOS DE PELÍCULAS

Noel Sarrió

Máster en Análisis de Datos 2024-2025

Índice

Resumen Ejecutivo.....	2
Introducción	4
Metodología.....	5
Análisis de Consultas SQL para la Base de Datos de Películas.....	6
Conclusiones y Recomendaciones.....	8

Título del Proyecto: Análisis Base de Datos de Películas

Autor: Noel Sarrió Ruiz

Fecha: 28 febrero 2025

Resumen Ejecutivo

Este informe presenta un conjunto de consultas SQL diseñadas para analizar una base de datos de películas, actores, alquileres y clientes. Las consultas cubren una amplia gama de escenarios comunes en la gestión de datos de una plataforma de alquiler de películas, proporcionando información clave sobre el comportamiento de los clientes, las películas más populares, los ingresos generados y la distribución de las películas en categorías.

El análisis incluye las siguientes áreas clave:

1. **Información sobre Películas:** Se realiza un análisis sobre las películas basándose en su clasificación, duración, categoría y frecuencia de alquiler. También se examinan películas con características específicas, como aquellas con una duración superior a 180 minutos o con una duración igual a la de una película específica.
2. **Actores:** Se exploran los actores según diferentes criterios, incluyendo su participación en películas, su número total de películas en las que han actuado y su participación en categorías específicas, como 'Sci-Fi' o 'Action'. También se identifican actores que no han participado en ninguna película.
3. **Clientes y Alquileres:** Se analizan los alquileres realizados por los clientes, identificando aquellos que han alquilado un mayor número de películas o generado mayores ingresos para la empresa. Se incluyen también las estadísticas sobre la duración de los alquileres y el número total de alquileres realizados por cliente.
4. **Estadísticas Agregadas y Cálculos Financieros:** El informe incluye el cálculo de estadísticas clave, como el promedio de ingresos generados, la desviación estándar de los pagos, y el número total de alquileres por mes. Estas estadísticas son fundamentales para comprender el rendimiento general de la plataforma.

5. **Consultas Avanzadas:** Algunas consultas incluyen la creación de vistas y tablas temporales, lo que permite almacenar resultados intermedios para análisis adicionales. También se analiza la viabilidad del uso de un CrossJoin, que no tiene valor práctico en este caso debido a la falta de relaciones lógicas entre las tablas involucradas.

Este conjunto de consultas SQL no solo proporciona una visión detallada sobre las operaciones de la plataforma, sino que también ofrece una base sólida para la toma de decisiones estratégicas, optimizando los recursos y mejorando la experiencia tanto para los clientes como para los operadores del sistema.

Introducción

El análisis de datos es una herramienta fundamental en la toma de decisiones empresariales, especialmente en sectores como el alquiler de películas, donde la gestión eficiente de la información es clave para mejorar el rendimiento y optimizar las operaciones. Este informe tiene como objetivo presentar un conjunto de consultas SQL diseñadas para analizar una base de datos de una plataforma de alquiler de películas. Las consultas cubren diversas áreas relacionadas con las películas, actores, clientes, alquileres y las estadísticas de pagos, con el fin de proporcionar una visión integral del negocio y sus operaciones.

En el entorno actual de negocios, los datos se han convertido en un activo estratégico. Las consultas SQL presentadas en este informe permiten extraer, filtrar y analizar información relevante que puede guiar decisiones sobre la gestión de inventarios, la estrategia de precios, las campañas de marketing y la optimización de la experiencia del cliente. Este análisis está diseñado no solo para profundizar en la comprensión de la base de datos, sino también para identificar oportunidades de mejora en los procesos de alquiler y distribución de películas.

A lo largo de este informe, se explorarán diferentes consultas que abordan aspectos clave del negocio, como la clasificación de películas, la frecuencia de alquileres, la duración de los alquileres, el comportamiento de los clientes y las estadísticas financieras, entre otros. Con este enfoque, se pretende ofrecer un análisis exhaustivo que pueda ser utilizado como base para futuras estrategias comerciales y operativas dentro de la plataforma de alquiler de películas.

Metodología

El análisis presentado en este informe se ha basado en el uso de consultas SQL aplicadas a una base de datos estructurada en un sistema de gestión de alquiler de películas. La metodología seguida se ha centrado en la extracción, manipulación y análisis de los datos disponibles, con el objetivo de proporcionar insights clave sobre las operaciones de la plataforma. A continuación, se detallan los principales pasos y enfoques utilizados en el proceso de análisis:

1. Estructura de la Base de Datos

La base de datos de la plataforma de alquiler de películas está organizada en diversas tablas, que incluyen información sobre películas, actores, clientes, alquileres y pagos. Las tablas clave utilizadas en este análisis incluyen:

- **Film:** Contiene información sobre las películas, como el título, la duración, la clasificación, el idioma, etc.
- **Actor:** Contiene detalles sobre los actores, incluyendo su nombre y apellido.
- **Rental:** Registra los alquileres realizados, con información sobre fechas de alquiler y devolución.
- **Payment:** Almacena los pagos realizados por los clientes durante los alquileres.

2. Extracción de Datos

El proceso comenzó con la extracción de datos relevantes de las tablas mencionadas. Se utilizaron consultas SELECT para obtener registros específicos y realizar filtrados según criterios como clasificación de las películas, duraciones, y actores involucrados. Se emplearon funciones de agregación como COUNT(), SUM(), AVG(), entre otras, para obtener métricas clave.

3. Análisis Descriptivo

Se realizaron consultas orientadas a obtener estadísticas descriptivas, como el total de ingresos generados por la plataforma, la distribución de alquileres por cliente y categoría de película, así como la identificación de películas y actores más frecuentes. Las consultas fueron estructuradas para analizar la base de datos de manera eficiente, respondiendo a preguntas como:

- ¿Cuáles son las películas más alquiladas?
- ¿Qué actores han participado en más películas?
- ¿Cuál es el comportamiento de los clientes en términos de alquiler y pago?

4. Filtrado y Agrupación de Datos

Se emplearon técnicas de filtrado para seleccionar datos específicos, tales como las películas con duración superior a un umbral determinado o aquellas con ciertas clasificaciones. Además, se utilizó el agrupamiento (GROUP BY) para realizar agregaciones por categorías, por ejemplo, calcular el número total de alquileres por cliente o por categoría de película.

5. Análisis de Tendencias Temporales

Se llevó a cabo un análisis temporal para identificar patrones y tendencias en los datos de alquiler, como la cantidad de alquileres por mes, la duración promedio de los alquileres, y las películas más alquiladas en determinados períodos.

6. Creación de Vistas y Tablas Temporales

Para optimizar el proceso de análisis, se crearon vistas y tablas temporales que permitieron almacenar resultados intermedios de las consultas más complejas. Esto facilitó el análisis incremental y la obtención de insights más detallados a partir de conjuntos de datos reducidos.

7. Validación y Refinamiento

Una vez obtenidos los primeros resultados, se revisaron las consultas para asegurarse de que los datos extraídos fueran correctos y reflejaran adecuadamente la estructura de la base de datos. Se realizaron ajustes en las consultas y se aplicaron filtros adicionales para mejorar la precisión de los resultados.

Análisis de Consultas SQL para la Base de Datos de Películas

Este conjunto de consultas SQL abarca una variedad de escenarios que pueden encontrarse en una base de datos de alquiler de películas. Las consultas están diseñadas para explorar diferentes aspectos de los datos, como la información de las películas, los actores, las categorías, los alquileres y los clientes. Por ejemplo:

Consultas sobre Películas y Clasificación

- **Películas con clasificación 'R':** Recupera todos los títulos de películas con una clasificación de 'R'.
- **Películas con duración mayor a 180 minutos:** Filtra las películas con una duración superior a 180 minutos.

- **Películas por categoría:** Filtra las películas que pertenecen a una categoría específica como 'Action' o 'Comedy'.
- **Películas alquiladas más de 8 días:** Identifica las películas que fueron alquiladas por más de 8 días.
- **Películas con duración igual a 'Dancing Fever':** Encuentra películas que tengan la misma duración que una película específica.

Consultas sobre Actores

- **Actores con actor_id entre 30 y 40:** Selecciona actores dentro de un rango de ID.
- **Actores con apellido 'Allen':** Filtra actores cuyo apellido sea 'Allen'.
- **Actores que han participado en películas de 'Sci-Fi':** Muestra los actores que han trabajado en películas de la categoría 'Sci-Fi'.
- **Actores sin películas:** Encuentra los actores que no han participado en ninguna película.
- **Actores que han trabajado en más de 40 películas:** Filtra actores que han actuado en más de 40 películas.

Consultas sobre Clientes y Alquileres

- **Total de dinero generado por la empresa:** Calcula el total de ingresos a partir de los pagos realizados por los clientes.
- **Clientes con más alquileres:** Muestra los clientes que han realizado más alquileres.
- **Clientes que han alquilado al menos 7 películas:** Identifica clientes que han alquilado un número significativo de películas.
- **Alquileres por cliente y película:** Realiza una relación entre clientes y sus respectivos alquileres.

Consultas sobre Datos Agregados y Estadísticas

- **Promedio, desviación estándar y varianza del total pagado:** Calcula las estadísticas sobre el total de dinero pagado.
- **Número total de alquileres por mes:** Agrupa los alquileres por mes y año, permitiendo analizar tendencias temporales.
- **Categorías de películas con duración superior a 110 minutos:** Agrupa las películas por categoría y calcula la duración promedio.

Consultas Avanzadas

- **Cross Join entre películas y categorías:** Muestra todas las combinaciones posibles de películas y categorías, aunque no tiene un valor significativo para este caso en particular.
- **Crear vistas y tablas temporales:** Varias consultas incluyen la creación de vistas o tablas temporales, lo que permite almacenar resultados intermedios para análisis adicionales.

Conclusiones y Recomendaciones

Este conjunto de consultas SQL cubre una amplia gama de funciones que puedes utilizar para analizar y manipular datos de una base de datos de alquiler de películas. Las consultas ayudan en la gestión de datos, proporcionan información sobre los actores y clientes, y permiten realizar análisis estadísticos de los alquileres y las películas.

Sugerencias para la práctica

1. **Optimización de Consultas:** Algunas consultas, como el uso de Cross Join, pueden ser mejoradas utilizando un Inner Join o Left Join, dependiendo del contexto.
2. **Uso de Funciones de Agregación:** Es importante utilizar funciones como COUNT, AVG, SUM y GROUP BY para obtener información agregada de manera eficiente.
3. **Trabajo con Fechas:** Las consultas que incluyen operaciones con fechas, como la duración del alquiler o las fechas de pago, son esenciales para un análisis temporal y deben optimizarse correctamente.