## ****Informe de Análisis de Datos de Alquiler de Películas****

### **1. Arquitectura de Datos y Arquetipo de Aplicación**

**Arquitectura de Datos:** Basándonos en las tablas proporcionadas, podemos inferir una arquitectura de datos relacional. Cada tabla representa una entidad específica:

* **film:** Información detallada sobre las películas, como duración, clasificación, características especiales.
* **inventory:** Registra las copias disponibles de cada película en cada tienda.
* **rental:** Detalla los alquileres realizados, incluyendo la fecha de alquiler y devolución.
* **customer:** Contiene información sobre los clientes, como nombre, dirección y estado de la cuenta.
* **store:** Informa sobre las tiendas, incluyendo el empleado encargado.

Las relaciones entre estas tablas son las siguientes:

* **film** y **inventory:** Una película puede tener múltiples copias en el inventario.
* **rental** y **inventory:** Cada alquiler se asocia con una copia específica del inventario.
* **rental** y **customer:** Cada alquiler está vinculado a un cliente.
* **store** y **inventory:** Las copias de las películas están ubicadas en una tienda específica.
* **store** y **customer:** Los clientes están asociados a una tienda específica.

**Arquetipo de Aplicación:** Este conjunto de datos parece ser parte de un sistema de gestión de videoclubes o una plataforma de streaming. La aplicación subyacente probablemente sea un sistema de información transaccional (SIT) que registra las operaciones diarias de alquiler de películas.

### **2. Análisis Exploratorio de Datos (EDA)**

**Observaciones Iniciales:**

* **Completitud de los datos:** Las tablas parecen tener una estructura consistente y sin valores nulos evidentes en los datos proporcionados.
* **Tipos de datos:** Los datos son una mezcla de numéricos (por ejemplo, rental\_duration, rental\_rate) y categóricos (por ejemplo, rating, special\_features).
* **Relaciones:** Las relaciones entre las tablas son claras y permiten realizar análisis detallados sobre los patrones de alquiler.

**Preguntas de Investigación Iniciales:**

* ¿Cuáles son las películas más populares (basadas en el número de alquileres)?
* ¿Cuál es la duración promedio de alquiler?
* ¿Cuál es el género de película más alquilado?
* ¿Existe una relación entre la calificación de una película y su popularidad?
* ¿Cuáles son los clientes más activos?

### **3. Preguntas de Negocio y Respuestas**

**Pregunta 1: ¿Cuál es la película más alquilada y cuál es la menos alquilada?**

* **Respuesta:** Para responder a esta pregunta, necesitaríamos unir las tablas rental y film por el campo film\_id. Luego, agrupar por film\_title y contar el número de alquileres. La película con el mayor número de alquileres sería la más popular.

**Pregunta 2: ¿Cuál es el género de película más rentable?**

* **Respuesta:** Suponiendo que tenemos un campo genre en la tabla film, podríamos calcular el ingreso total por género multiplicando la tarifa de alquiler por el número de alquileres por género.

**Pregunta 3: ¿Cuál es la duración promedio de alquiler de las películas clasificadas como 'G'?**

* **Respuesta:** Filtraríamos los datos de la tabla rental por películas clasificadas como 'G' y calcularíamos la duración promedio de los alquileres.

**Pregunta 4: ¿Existe una correlación entre la duración de una película y su tarifa de alquiler?**

* **Respuesta:** Calcularíamos el coeficiente de correlación de Pearson entre las columnas length y rental\_rate de la tabla film para determinar si existe una relación lineal entre ambas variables.

**Pregunta 5: ¿Cuál es el cliente que ha generado más ingresos para el videoclub?**

* **Respuesta:** Uniríamos las tablas rental y customer por customer\_id. Luego, calcularíamos el ingreso total generado por cada cliente y encontraríamos al cliente con el ingreso más alto.

### **4. Conclusiones**

El conjunto de datos proporcionado ofrece una base sólida para analizar el comportamiento de alquiler de películas. Al realizar un análisis más profundo, podemos obtener insights valiosos como:

* **Popularidad de películas:** Identificar las películas más y menos populares.
* **Preferencias de los clientes:** Analizar los géneros y las clasificaciones más demandadas.
* **Rendimiento de las tiendas:** Comparar el número de alquileres y los ingresos por tienda.
* **Efectividad de las estrategias de precios:** Evaluar la relación entre la tarifa de alquiler y el número de alquileres.

**Siguientes paso con el ETL:**

* **Limpieza de datos:** Identificar y corregir cualquier inconsistencia o error en los datos.
* **Análisis más detallado:** Utilizar técnicas estadísticas y de visualización para explorar los datos en profundidad.
* **Modelado predictivo:** Crear modelos para predecir la demanda de películas o el comportamiento de los clientes.

**Recomendaciones:**

* **Incorporar más datos:** Agregar datos sobre las características de los clientes (edad, género) o información sobre las promociones podría enriquecer el análisis.
* **Visualización:** Utilizar herramientas de visualización como Tableau o Power BI para crear gráficos y dashboards que faciliten la comprensión de los resultados.
* **Análisis de series temporales:** Analizar la evolución de los alquileres en el tiempo para identificar tendencias y patrones estacionales.

Este informe presenta un análisis preliminar de los datos. Un análisis más exhaustivo permitiría tomar decisiones más informadas sobre la gestión del videoclub y la estrategia de negocio.

**Nota:** Para realizar este análisis, sería necesario utilizar una herramienta de análisis de datos como Python (con librerías como Pandas y NumPy) o un software de Business Intelligence.