Alquiler de películas para la datawarehouse

Maikol Steven pelaéz Betancur

Universidad Uninpahu

Ing . Software

22 de abril 2019

**Alcance:**

En este proyecto se justificara la creación de un datawarehouse. Los datos estarán almacenados históricamente, conjuntamente lo propuesto en este proyecto facilitara el procesamiento y consultas obteniendo información valiosa para el negocio.

Este proyecto se centrara en el diseño e implementación de datawarehouse y las acciones ETL con 5 fuentes proporcionadas con información inconsistente para realizar la actividad de crear reglas de negocio para su extracción trasformación y carga a la datawarehouse.

**Objetivo:**

Realizar una datawarehouse que cumpla con los estándares de las tiendas de alquiler de películas al analizar la información obtenida para generar toma de decisiones para la compra de artículos para el alquiler y así disparar el alquiler de los productos propuestos en el proyecto y así fidelizar a los clientes a las tiendas.

**Objetivos específicos:**

- obtener información valiosa para anticipar escases de películas más solicitados por el cliente.

- realizar consultas ordenadas que permita analizar nuevas estrategias comparando la periodicidad día, mes año alquiladas, solicitadas por los clientes.

-centralizar la información de las tiendas para definir nuevas estrategias de negocio se quiere analizar el alquiler realizado para cada tienda y sus ganancias para así estudiar los puntos de alquiler que beneficios aportan.

-obteniendo la información de los clientes fieles que adquieren más alquiler para darles promociones y fidelizarlos a las tiendas.

**Variables para las 5 fuentes:**

|  |  |
| --- | --- |
| fecha alquiler | fecha |
| fecha devolución | fecha |
| estado alquiler | Número |
| Identificación cliente | Número |
| nombre cliente | texto |
| apellido cliente | texto |
| dirección cliente | texto |
| categoría película | texto |
| Nombre película | texto |
| descripción película | texto |
| año de estreno | fecha |
| duración película | numero |
| Costo alquiler | numero |

**Transformaciones**

**Regla 1**

* La tienda requiere la fecha en formato año/mes/dia para realizar las consultas de películas más alquiladas.

**Transformación:**

02-01-2015 00:00:00 🡪 1995/02/15 00:00:00

**Regla 2**

* Se debe eliminar los clientes duplicados de las fuentes proporcionadas para tener datos totalmente integrales para realizar un proceso de negocio con estadísticas reales y más exactas posibles.

**Transformación**

|  |  |  |  |
| --- | --- | --- | --- |
| id | 10154548458 | nombre | pepito |
| id | 10154548458 | nombre | pepito |

Validación de la etl de duplicados para la carga sin duplicados

|  |  |  |  |
| --- | --- | --- | --- |
| id | 10154548458 | nombre | pepito |

**Regla 3**

* En un campo especifico concatenar el nombre y apellido para tener un reporte de un solo renglón de los nombres del cliente.

**Transformación**

Nombre: pepito

Apellido: Perez ----> Nombres: pepito Perez

**Regla 4**

* Validar las fuentes que proporcionan no tengan valores nulos en los campos definidos para la datawarehouse como fechas, nombres de clientes, nombre de película, categoría y costo alquiler si alguno de los campos esta nulo no realizar la carga de la información a la datawarehouse si por algún motivo el resto de campos es nullo autocompletarlo con “ **NO aplica ”**

**Transformación**

|  |  |
| --- | --- |
| fecha alquiler | null |
| fecha devolución | 10/02/2006 00:00:00 |
| nombre | pepito |
| dirección | Calle 3 |

|  |  |
| --- | --- |
| fecha alquiler | 10/02/2006 12:30:00 |
| fecha devolución | 10/02/2006 17:00:00 |
| nombre | pepito |
| dirección | null |

Aplicando la regla queda de esta forma la primera información no cumple no se agrega para efectuar la carga

|  |  |
| --- | --- |
| fecha alquiler | 2006/02/10 12:30:00 |
| fecha devolución | 2006/02/10 17:00:00 |
| nombre | pepito |
| dirección | No aplica |

Regla 5

Diseño:

