

Netflix Data Warehouse

Susan Hidalgo, Brenda Rojas, Cristopher Rodriguez, Melani Vargas

shidalgo70985@ufide.ac.cr

brojas60567@ufide.ac.cr

crodriguez00279@ufide.ac.cr

mvgargas30424@ufide.ac.cr

Universidad Fidélitas, Ingeniería en
Sistemas Computacionales, Heredia,
Costa Rica

Data Warehouse

(SC-602)

Grupo #1

Resumen– Se quiere realizar una investigación a profundidad de la empresa Netflix para poder llegar a crear un Data Warehouse que logre satisfacer a la empresa y sus necesidades, tomando en cuenta datos proporcionados por el sitio Kaggle, donde se obtuvo información de los títulos que actualmente presenta la aplicaciones e información como sus géneros, fechas de lanzamiento, países entre otros.

Palabras Claves– Data Warehouse, Netflix, Kaggle, Streaming

Abstract– We want to carry out an in-depth investigation of the Netflix company in order to create a Data Warehouse that manages to satisfy the company and its needs, taking into account data provided by the Kaggle site, where information was obtained on the titles that the applications currently present and information such as their genres, release dates, countries among others.

Keywords– Data Warehouse, Netflix, Kaggle, Streaming

I. CAPÍTULO DE INTRODUCCIÓN

➤ INTRODUCCIÓN

Para nuestro proyecto de investigación del curso Data Warehouse hemos decidido basarnos en la famosa aplicación Netflix, la cual se basa en presentarle al comprador de la mensualidad de la aplicación una amplia gama de películas, series y shows televisivos donde se pueden descargar para ver sin internet, ver las películas o series en múltiples idiomas, entre otras muchísimas ventajas que nos ofrece el sitio.

En sus inicios Netflix era un centro de alquiler y venta de DVDs, pero sus fundadores Reed Hastings y Marc Randolph vieron un gran potencial en la venta de películas en línea, su nombre se debe a la unión de las palabras “internet” y “flicks” lo cual significa películas

Con el pasar de los años Netflix ha ganado un puesto en el mercado muy importante siendo prácticamente la aplicación de

películas y series número 1 a nivel mundial, esto permitiéndoles no solo ser una plataforma de streaming sino que también les dio la oportunidad de empezar a grabar su propio contenido incluyendo series y películas, esto se debe también gracias a la pandemia ya que en el 2020 su uso se disparó de una manera estratosférica, actualmente cuenta con una base de datos muy amplia con diferentes países, fechas de lanzamientos, géneros de películas y series, lo que nos permite poder desarrollar diferentes gráficas, esquemas e investigaciones para poder profundizar más en una funcionalidad que logre satisfacer a la empresa.

➤ OBJETIVOS

Objetivo general:

Desarrollar un Data Warehouse funcional que permita la optimización de un análisis estratégico de contenido en Netflix mediante la integración eficiente de datos proveniente de diversas fuentes, contando el sistema con tiempos de respuesta inferiores a un segundo para consultas claves y ofrecer soporte a la toma de decisiones empresariales relacionadas con la planificación de contenido, análisis de popularidad y tendencias del mercado.

Objetivos específicos:

- Implementar un proceso de extracción, transformación y carga (ETL) que procese datos clave, como géneros de contenido, países, métricas de popularidad y reseñas de usuario para garantizar la calidad, consistencia y disponibilidad de información en el Data Warehouse..
- Diseñar un modelo multidimensional en el Data Warehouse basado en un modelo estrella para analizar el comportamiento del consumo de contenido según criterios como género, país, fecha de lanzamiento y popularidad, optimizando así el diseño para consultas rápidas y análisis detallados.
- Integrar herramientas de virtualización como Power BI, para generar reportes interactivos que muestran distintos

patrones de consumo, tendencias de popularidad por género y país, y métricas de desempeño financiero, ayudando a que las visualizaciones tomen decisiones estratégicas basadas en datos.

➤ ANTECEDENTES

Al pasar los años la tecnología ha sido parte de cada innovación que se ha presentado en el mundo, esto también fue visto en el alquiler o renta de películas, ya que la fundación de Netflix se remonta a que uno de sus fundadores se atrasó en devolver una película a la tienda de videos y se le cobró una multa, debido a esto él decidió crear un sistema nuevo donde se pagará por mes y no existieran esos cobros o multas por atrasos de devolución, al pasar los años la venta de películas en línea a sido todo un éxito permitiendo que la empresa de Netflix creciera y se convirtiera en lo que hoy en día conocemos, tantos años de trayectoria involucra muchísimos datos e información la cual podemos usar para dar el mejor funcionamiento y una experiencia de la más alta calidad a todos los usuarios.

➤ JUSTIFICACIÓN

El dataset de *Netflix Movies and TV shows* que se encuentra disponible en la página Kaggle, proporciona valiosa información acerca de los títulos que actualmente son más vistos en la plataforma, incluyendo detalles como género, país, fechas de lanzamiento, entre otros. Dado al gran volumen e importancia de estos datos para la toma de decisiones estratégicas en la industria del entretenimiento, se plantea el desarrollo de un Data Warehouse funcional para el análisis y optimización del catálogo de contenido. A su vez, la solución de este proyecto facilitará la integración de diversas fuentes de datos y ayudará con la identificación de tendencias enfocadas en el consumo, popularidad de géneros y análisis de mercado, además la creación de un Data Warehouse optimizará la organización y acceso de datos, mejorando la eficiencia en la toma de decisiones, como la planificación de contenido basado en predicciones de futuras demandas.

REQUERIMIENTOS FUNCIONALES

- Fuentes de datos: Se utilizarán cinco datasets relacionados con Netflix, disponibles en la plataforma Kaggle. Estos incluyen datos sobre películas y series, reseñas de usuarios, comportamiento bursátil, y la base de usuarios de Netflix. Los datos estarán en formato CSV y serán transformados mediante Pentaho antes de ser almacenados en el Data Warehouse.
- Componentes ETL: El proceso ETL será gestionado por Pentaho, el cual se encargará de extraer los datos desde los archivos CSV, transformarlos y cargarlos en el SQL Server.

CARACTERÍSTICAS DE LA SOLUCIÓN

- Pentaho: Una herramienta que se utilizará para llevar a cabo los procesos ETL, es decir, extracción, transformación y carga. Se encargará de integrar los datasets encontrados,, limpiando y transformando los datos según sea necesario antes de almacenarlos en el Data Warehouse.
- SQL Server: Será el sistema de gestión de bases de datos

donde se implementará el Data Warehouse. Aquí se crearán las tablas de hechos y dimensiones, organizadas bajo un modelo multidimensional como el esquema estrella, lo que permitirá almacenar y organizar la información de forma eficiente.

- Power BI: Se utilizará para generar reportes y visualizaciones a partir de los datos almacenados en el Data Warehouse, de modo que se pueda analizar el rendimiento de los contenidos de Netflix, identificar tendencias y realizar seguimientos de datos financieros.

II. DIAGRAMAS DE ENTIDAD-RELACIÓN DEL MODELADO DIMENSIONAL

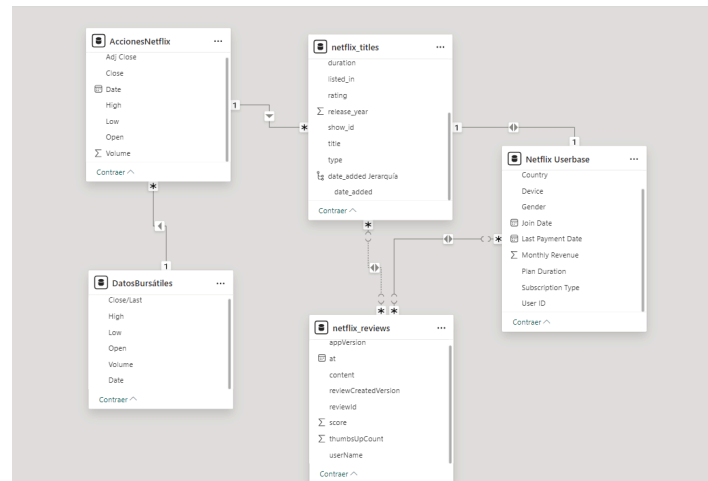


Fig 1 Modelado de Base de Datos de PowerBI

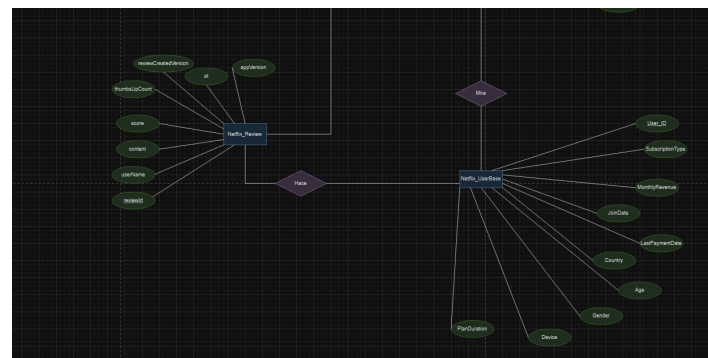


Fig 2 Diagrama de Entidad relación draw.io

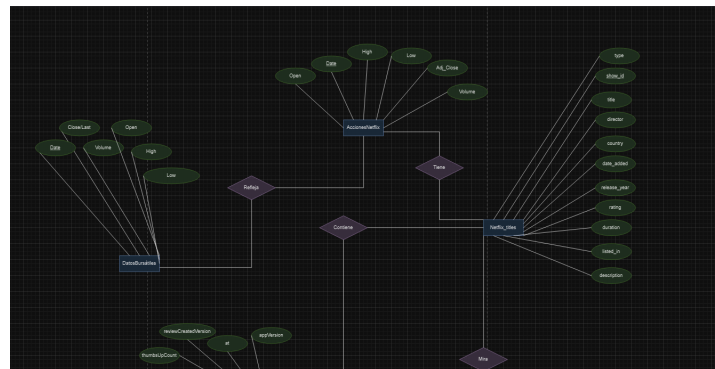


Fig 3 Diagrama de Entidad relación draw.io

https://drive.google.com/file/d/1x2qV4JiYZA8nhZ5y6_t5Vxdqmzinhg/view?usp=sharing

III. DEFINICIÓN DE LAS FUENTES DE DATOS

Los datos utilizados en este proyecto provienen de la plataforma "Kaggle", que ofrece miles de datasets gratuitos para la práctica de análisis de datos. En nuestro caso, decidimos trabajar con datos relacionados con Netflix, una de las plataformas de streaming de películas y series más reconocidas a nivel mundial, con millones de usuarios. Esto nos permitió acceder a una amplia cantidad de información disponible para el análisis.

Los archivos de datos que utilizamos en el proyecto son:

1. **NETFLIX Stock Data**
2. **Netflix Reviews [DAILY UPDATED]**
3. **Netflix Movies and TV Shows**
4. **Datos Bursátiles de Netflix**
5. **Netflix User Base**

Estos archivos contienen una variedad de información, como reseñas de usuarios, rankings de películas y series, los países donde Netflix tiene mayor audiencia, y datos sobre el comportamiento bursátil de la compañía. Todos estos datasets están en formato CSV (texto plano) y fueron abiertos en Excel para obtener una vista más clara y ordenada de su contenido.

El siguiente paso fue integrar los datos en Power BI, donde construimos un modelo de base de datos relacional. A partir de este modelo, generamos un diagrama entidad-relación (ERD) que representaba las conexiones entre las diferentes tablas y los datos.

En total, utilizamos cinco datasets con información diversa. Los datos se organizaron en tablas separadas, estableciendo las relaciones correspondientes a través de claves primarias (PK) y claves foráneas (FK), lo que nos permitió conectar adecuadamente los diferentes conjuntos de datos y realizar un análisis más profundo y visual de la información.

IV. FUENTES DE DATOS

Netflix es una de las empresas de streaming más importantes a nivel mundial, es conocida por su amplio repertorio de entretenimiento visual, la compañía se apoya en una cantidad casi inimaginable de datos en sus bases de datos, las cuales son usadas para estrategias comerciales y para mejorar las experiencias de los usuarios, al igual que para tener un control de su plataforma y

quienes tengan acceso a ella.

Para nuestro proyecto se estarán analizando diversas fuentes de datos en archivos CSV, las cuales reflejan diferentes áreas importantes de la compañía, a continuación se describirán 5 archivos CSV con datos de la empresa y sus explicaciones de la importancia de cada uno.

- **AccionesNetflix**

Este archivo contiene los datos históricos de los precios de las acciones de la compañía de Netflix, cuenta con columnas como Date la cual nos indica la fecha en que los valores fueron registrados, Open indica el precio con el que se abrió la acción en la fecha indicada, High indica el valor más alto al que llegó la acción en la fecha indicada, Low indica el valor más bajo al que llegó la acción en la fecha indicada, Close nos indica en qué momento se cerró las acciones de la fecha indicada, Adj Close indica el precio del cierre ajustado y por último tenemos Volume que es el número de acciones negociadas durante la fecha indicada.

Las columnas tienen valores de tipo numérico decimal o entero y de fecha, los datos comienzan el 23 de mayo del 2002 y luego van incrementando día con día.

El uso que representan estos datos en nuestro proyecto es poder analizar cómo va cambiando el valor en la bolsa de acciones de la empresa de Netflix.

- **Datos Bursátiles**

Esta tabla básicamente contiene los datos de las acciones en el mercado financiero del año 2020, las columnas representan los precios y volúmenes de las acciones por día, cuenta con una columna Date la cual tiene el formato de mes/ día/ año, eso se usa para poder analizar los patrones en los datos, luego contiene un Close/Last el cual se encarga de mostrar el precio del cierre de la acción en ese día específico, luego está el Volume que indica el número de acciones totales en el día indicado, el Open indica el monto con el que inició la acción en ese día, High nos indica el monto más alto de la acción en ese día específico y por último esta el Low que

Este CSV lo incorporamos para poder tener los datos de la parte financiera tanto antiguos los cuales son los del CSV anteriores como los datos un poco más actualizados y así poder comparar los cambios a través del tiempo.

Este CSV contiene la información de los usuarios, contiene datos como, su User ID para poder llevar un control ordenado, la Subscription Type ya sea el plan básico, el estándar o el premium, la diferencia de un plan a otro es que en el básico cuenta con características limitadas, en el estándar se puede tener acceso a múltiples pantallas y calidad HD y en el premium se tiene acceso ilimitado a las pantallas y calidad UHD. También se tiene los datos de el Monthly Revenue el cual es el ingreso mensual generado por cada usuario, es decir lo que gasta cada uno según el tipo de suscripción que tengan, también tiene la fecha en la que se unió el usuario al servicio, la fecha del ultimo día que el usuario realizó el pago, el país, la edad, el género el dispositivo que utiliza el usuarios es decir si un celular, computadora o televisor y por último también se tiene la duración del plan, si paga por mes o por año.

- Reviews

Este CSV es de suma importancia para poder ver las preferencias de los usuarios y sus opiniones tanto de la plataforma como de los títulos que la misma presenta, ya sea

- Titles

Estos datos son importantes porque se pueden agrupar los títulos por año, género, duración, director, actores, entre otros. Esto permite tener un orden de búsqueda y de repertorio lo más acomodado y actualizado posible, esto con el fin de tener la información detallada.

The diagram illustrates a data lake schema with the following tables and their attributes:

- netflix_titles**
 - show_id
 - type
 - title
 - director
 - cast
 - country
 - date_added
 - release_year
 - rating
 - duration
- Netflix_Reviews**
 - review_id
 - user_id
 - content
 - score
 - thumbs_up_count
 - review_created_datetime
 - all
 - appVersion
 - Repetitions
- HechosNetflix**
 - Fecha
 - UserID
 - show_id
 - review_id
 - DatosBursatiles
 - AccionesNetflix
 - Reproducciones
 - TemporoReproduccion
 - ImpresioManual
- DatosBursatiles**
 - Id
 - Fecha
 - Close
 - Volume
 - Open
 - High
- AccionesNetflix**
 - Id
 - Fecha
 - Open
 - High
 - Low
 - Close
 - Avg(Close
 - Volume
- Netflix Userbase**
 - UserID
 - SubscriptionType
 - MonthlyRevenue
 - JoinDate
 - LastPaymentDate
 - Country
 - Age
 - Gender
 - Device
 - PlanDuration

Relationships are indicated by lines connecting the tables:

- netflix_titles** is connected to **HechosNetflix** via a line from `show_id` to `show_id`.
- Netflix_Reviews** is connected to **HechosNetflix** via a line from `review_id` to `review_id`.
- HechosNetflix** is connected to **DatosBursatiles** via a line from `DatosBursatiles` to `Id`.
- HechosNetflix** is connected to **AccionesNetflix** via a line from `AccionesNetflix` to `Id`.
- Netflix Userbase** is connected to **HechosNetflix** via a line from `UserID` to `UserID`.

Fig 4 Modelado Dimensional SOL SERVER

VI. PROCESOS ETL

Flujo para insertar los datos desde el csv de reviews en la tabla del schema SQL server. Se utilizó el paso Sort Rows para ordenar los datos de manera ascendente y que de este modo, el paso Unique Rows pueda filtrar las columnas con Pks repetidas. También se utilizó Filter rows para filtrar los campos con datos en null.

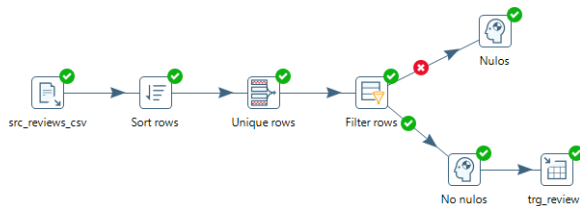


Fig.5 Flujo de la tabla Reviews

Flujo para transportar los datos desde el csv de datos bursátiles hasta la tabla en el SQL Server



Fig.6 Flujo de la tabla de datos bursátiles

Flujo para insertar los datos del csv de acciones de Netflix en su tabla correspondiente en el SQL Server. Se utilizó el paso Select Values para cambiar el tipo de dato de algunas columnas en el csv y que de este modo, no existiera algún tipo de conflicto a la hora de hacer la inserción.

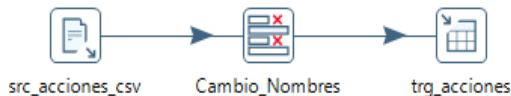


Fig.7 Flujo de la tabla Acciones

Para el flujo del csv de los títulos de películas se utilizó nuevamente el paso Filter Rows para filtrar las columnas con datos nulos.

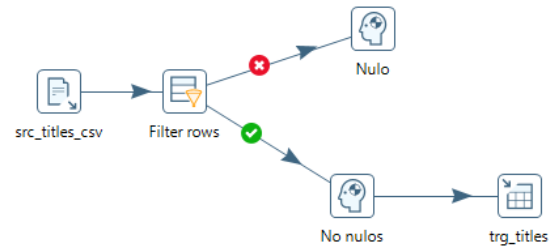


Fig.8 Flujo de la tabla Titles

Finalmente para la transformación de la Userbase de Netflix se utilizó el paso Select Values para cambiar el tipo de dato de algunas columnas, igualmente para evitar conflictos con la inserción.

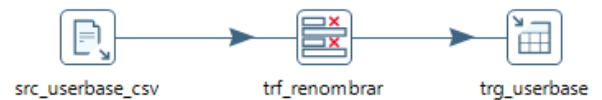


Fig.9 Flujo de la Userbase de Netflix

VII.

AVANCE DE LOS METADATOS

```
CREATE DATABASE Grupo1;
GO
```

```
Drop database Grupo1;
```

```
USE Grupo1;
GO
```

--Creación de las tablas

```
CREATE TABLE DatosBursatiles (
    Id INT PRIMARY KEY IDENTITY(1,1),
    Fecha DATE NOT NULL,
    Close_ DECIMAL(10,2),
    Volume INT,
    Open_ DECIMAL(10,2),
    High DECIMAL(10,2),
    Low DECIMAL(10,2)
);
```

```
CREATE TABLE AccionesNetflix (
    Id INT PRIMARY KEY IDENTITY(1,1),
    Fecha DATE NOT NULL,
    Open_ FLOAT,
    High FLOAT,
```

```

Low FLOAT,
Close_ FLOAT,
Adj_Close FLOAT,
Volume INT
);

CREATE TABLE Netflix_Reviews (
  reviewId VARCHAR(255) PRIMARY KEY,
  userName NVARCHAR(255),
  content TEXT,
  score INT,
  thumbsUpCount INT,
  reviewCreatedVersion VARCHAR(50),
  at DATE,
  appVersion VARCHAR(50)
);

ALTER TABLE Netflix_Reviews
ADD Repetidos INT;

```

```

CREATE TABLE netflix_titles (
  show_id VARCHAR(10) PRIMARY KEY,
  type VARCHAR(50),
  title VARCHAR(255),
  director VARCHAR(255),
  cast TEXT,
  country VARCHAR(255),
  date_added DATE,
  release_year INT,
  rating VARCHAR(10),
  duration VARCHAR(50),
  listed_in VARCHAR(255),
  description TEXT
);

```

```

CREATE TABLE Netflix_Userbase (
  UserID INT PRIMARY KEY,
  SubscriptionType VARCHAR(50),
  MonthlyRevenue INT,
  JoinDate DATE,
  LastPaymentDate DATE,
  Country VARCHAR(100),
  Age INT,
  Gender VARCHAR(10),
  Device VARCHAR(50),
  PlanDuration VARCHAR(50)
);

```

```

CREATE TABLE HechosNetflix (
  Fecha DATE NOT NULL,
  UserID INT,
  show_id VARCHAR(10),
  reviewId VARCHAR(255),
  DatosBursatilesId INT,
  AccionesNetflixId INT,

```

```

Reproducciones INT,
TiempoReproducido DECIMAL(10,2),
IngresoMensual DECIMAL(10,2),

```

```
PRIMARY KEY (UserID, show_id),
```

```

FOREIGN KEY (UserID) REFERENCES
Netflix_Userbase(UserID),
FOREIGN KEY (show_id) REFERENCES
netflix_titles(show_id),
FOREIGN KEY (reviewId) REFERENCES
Netflix_Reviews(reviewId),
FOREIGN KEY (DatosBursatilesId)
REFERENCES DatosBursatiles(Id),
FOREIGN KEY (AccionesNetflixId)
REFERENCES AccionesNetflix(Id)
);

```

Drop table HechosNetflix

```

--Actualizar fechas para pruebas
UPDATE Netflix_Userbase
SET JoinDate = '2020-07-01', LastPaymentDate =
'2020-07-31'
WHERE JoinDate IS NULL OR JoinDate >
'2020-01-01';

```

```

INSERT INTO HechosNetflix (Fecha, UserID,
show_id, reviewId, DatosBursatilesId,
AccionesNetflixId, Reproducciones,
TiempoReproducido, IngresoMensual)
SELECT

```

```

  db.Fecha,
  nu.UserID,
  nt.show_id,
  nr.reviewId,
  db.Id AS DatosBursatilesId,
  an.Id AS AccionesNetflixId,
  COUNT(nt.show_id) AS Reproducciones,
  SUM(0.5 * nu.MonthlyRevenue) AS
TiempoReproducido,
  nu.MonthlyRevenue AS IngresoMensual
FROM
  DatosBursatiles db
LEFT JOIN
  AccionesNetflix an ON db.Fecha = an.Fecha
LEFT JOIN
  Netflix_Userbase nu ON db.Fecha BETWEEN
nu.JoinDate AND nu.LastPaymentDate
LEFT JOIN
  Netflix_Reviews nr ON nr.at = db.Fecha
LEFT JOIN
  netflix_titles nt ON nr.reviewId IS NOT NULL

```


GROUP BY

db.Fecha, nu.UserID, nt.show_id, nr.reviewId,
db.Id, an.Id, nu.MonthlyRevenue;

--Métricas de la tabla hechos

--Reproducciones por título y región

```
SELECT N.Country, T.title,
SUM(H.Reproducciones) AS TotalReproducciones
FROM HechosNetflix H
JOIN Netflix_Userbase N ON H.UserID =
N.UserID
JOIN netflix_titles T ON H.show_id = T.show_id
GROUP BY N.Country, T.title;
```

SELECT*FROM HechosNetflix;

REPORTES EN POWER BI

Recuento de AuditID, Recuento de RecordID y Suma de Version por Año

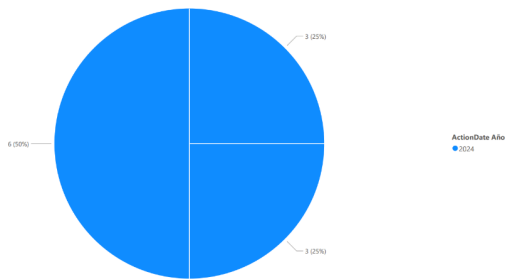


Fig 14 Auditoria Base de Datos



Fig 15 Resumen Acciones Netflix

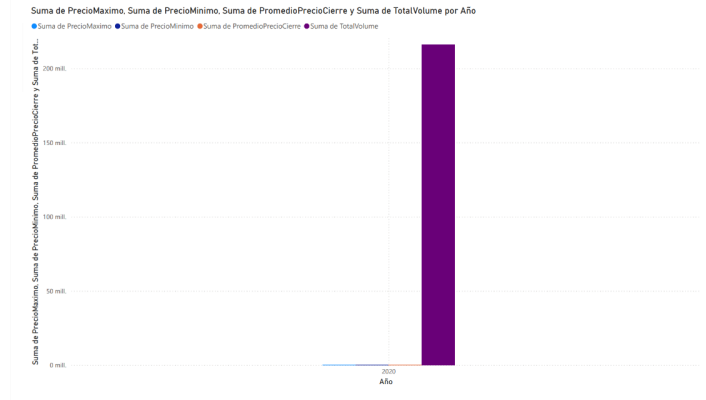


Fig 16 Resumen Datos Bursátiles

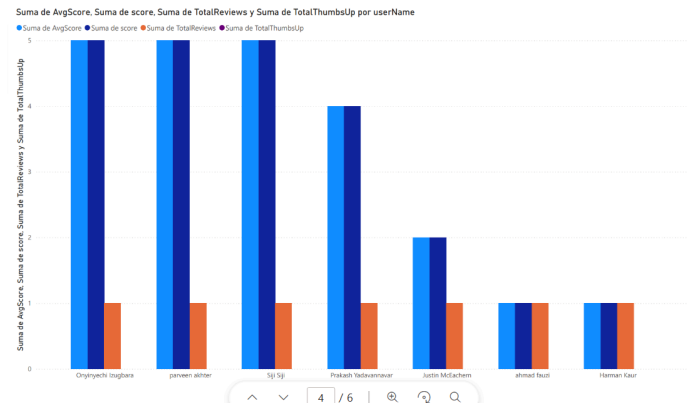


Fig 17 Resumen Netflix Reviews

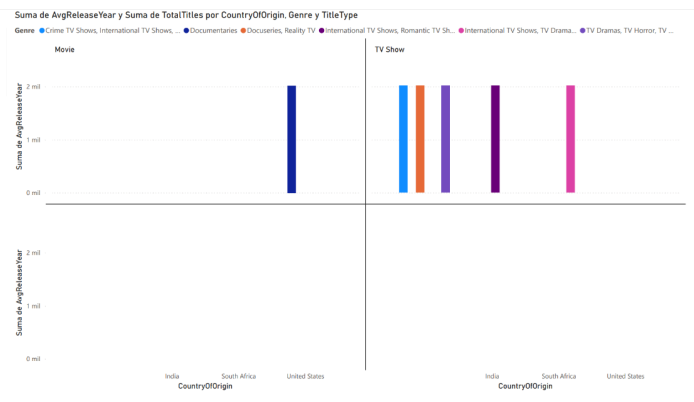


Fig 18 Resumen Netflix titles

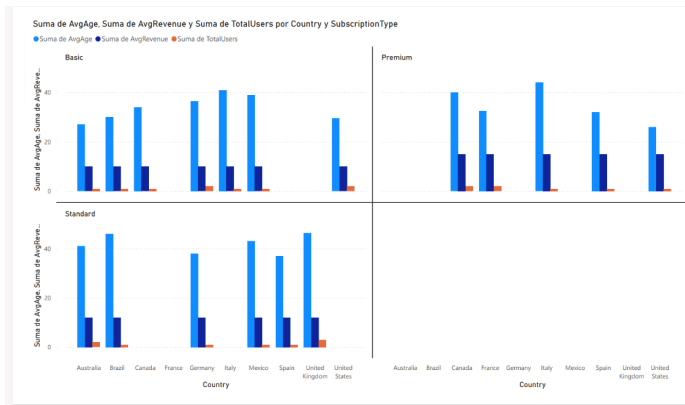


Fig 19 Resumen Netflix Userbase

[Reportes Netflix.pdf](#)

VIII. PROCESOS Y JUSTIFICACIÓN DE LAS DECISIONES TÉCNICAS

En esta parte se llevaron a cabo varios procesos tales como la extracción de los datos de múltiples fuentes, principalmente de archivos CSV de contenido, base de datos de reseñas y APIs con datos externos para saber las tendencias, el proceso de conexión con la fuente de los datos se hizo mediante SQL Management Studio, los datos se transformaron para normalizarlos con nombres que fueran aptos o entendidos tanto para SQL como para Pentaho, esto se hizo mediante el recurso de Pentaho, también se realizó carga de datos al Data Warehouse además de la generación de los reportes de los datos mediante Power BI que muestran las visualizaciones más relevantes.

Estas herramientas fueron utilizadas ya que nos parecieron que eran las más hábiles y útiles para poder llevar a cabo los procesos, Pentaho fue de gran ayuda para poder combinar datos de catalogo con métricas de usuarios en un solo flujo ETL, el uso de Power BI fue vital para poder llevar a cabo los gráficos mostrados de consumo mensuales, de género más popular, de tendencias entre otros.

El almacenamiento en SQL Management Studio fue la clave ya que se puede integrar de manera fácil tanto con Power BI como con Pentaho, la elección de métricas y dimensiones fue basada en dimensiones como país y fechas para detectar regiones que presenten patrones, esto para poder calificar el promedio de uso o analizar la popularidad del contenido, estas decisiones nos permiten responder incógnitas como ¿Cual generó fue el más visto en el 2020?

IX. RESULTADOS OBTENIDOS

Algunos de los resultados obtenidos fue la identificación

no sólo de género sino de regiones con mayor popularidad, se registraron las acciones registradas en fechas determinadas, algunos de los gráficos de Power BI presentados se agrupan por años para poder llevar un orden de crecimiento anual, también se registran las tendencias anuales, se muestra la suma del precio de las acciones, el precio de apertura de las acciones y también el volumen total de acciones negociadas.

También se logró registrar los datos de ciertos usuarios desde el año 2021 al 2023, contiene los tipos de datos de las cuentas, ya sean estándar, básico o premium, también lleva un registro de la fecha en la que se unieron, el último día de pago, año en el que realizaron su suscripción, esto con el fin de poder llevar la información importante del usuario controlada y poder análisis de los datos recopilados.

Además se tiene registro de todas las reseñas que se han hecho de las películas de la plataforma, con los datos de cuando se creó la reseña, aplicación o sitio web donde se escribió la reseña, las reseñas se tienen registradas desde el 2018 hasta la actualidad (2024).

Mediante esta recopilación y diferentes muestras de los datos podemos dar por concluidos los objetivos de nuestro proyecto, satisfactoriamente podemos decir que contamos con un Data Warehouse donde se tiene la suficiente cantidad de datos para poder hacer análisis a futuro o proyecciones de que le podría beneficiar más a la compañía de Netflix, se pueden ver reflejados desde los puntos débiles, hasta las mejores temporadas de la aplicación.

X. CONCLUSIONES

Se puede concluir con que los objetivos planteados en el inicio de nuestro proyecto se lograron cumplir con éxito, se logró implementar un Data Warehouse funcional que integra datos de múltiples fuentes, optimizando los procesos de análisis, esto permite que la generación de información valiosa sea tomada en decisiones importantes, como identificar tendencias, géneros, o países con mayor uso de la aplicación Netflix.

Se logró también la optimización del análisis de datos, al igual que el aporte del proceso ETL el cual es el proceso de extracción, transformación y carga la cual fue esencial para garantizar la calidad y disponibilidad de los datos en el Data Warehouse.

El impacto en la toma de decisiones lo podemos ver reflejado mediante nuestros reportes de Power BI, los cuales permiten identificar patrones claves.

Una recomendación a futuro que podemos dar es seguir ampliando el alcance del Data Warehouse, integrando datos como la interacción con redes sociales y métricas distintas con otros datos, para poder tener un panorama aún

más amplio del que estamos obteniendo actualmente.

XI. REFERENCIAS BIBLIOGRÁFICAS

- [1] A. González, “Ésta es la historia del origen de Netflix, la empresa que desplazó a Blockbuster”, *Grupo Milenio*, 29-ago-2023. [En línea]. Disponible en: <https://www.milenio.com/espectaculos/television/la-historia-detras-de-netflix-origenes-y-fundacion>. [Consultado: 15-oct-2024].

- [2] W. Guerra y M. R. Ichaso, “Cronología de Netflix: así se convirtió en el gigante del streaming”, *CNN en Español*, 18-jul-2022. [En línea]. Disponible en: <https://cnnespanol.cnn.com/2022/07/18/cronologia-netflix-asi-se-convirtio-gigante-streaming-orix/>. [Consultado: 15-oct-2024].

- [3] Abhishek. “NETFLIX Stock Data”. Kaggle: Your Machine Learning and Data Science Community. Accedido el 16 de octubre de 2024. [En línea]. Disponible: <https://www.kaggle.com/datasets/abhiram8/netflix-stock-data/code>

- [4] A. Kumar. “Netflix Reviews [DAILY UPDATED]”. Kaggle: Your Machine Learning and Data Science Community. Accedido el 16 de octubre de 2024. [En línea]. Disponible: <https://www.kaggle.com/datasets/ashishkumarak/netflix-reviews-playstore-daily-updated/data>

- [5] A. Kumar. “Netflix Reviews [DAILY UPDATED]”. Kaggle: Your Machine Learning and Data Science Community. Accedido el 16 de octubre de 2024. [En línea]. Disponible: <https://www.kaggle.com/datasets/ashishkumarak/netflix-reviews-playstore-daily-updated/data>