



SQL CASE STUDY
ON

NETFLIX

- By Sai Charan

Overview

- Introduction
- Objective & Scope
- Workflow
- Data Collection & Loading
- Data Cleaning
- Data Analysis
- Insights & Results

Introduction

This project focuses on analyzing Netflix data to uncover viewer behaviors and content.

By refining data types, resolving inconsistencies, and leveraging SQL queries, the analysis aims to provide actionable insights for optimizing content strategies and enhancing user experience on the platform.

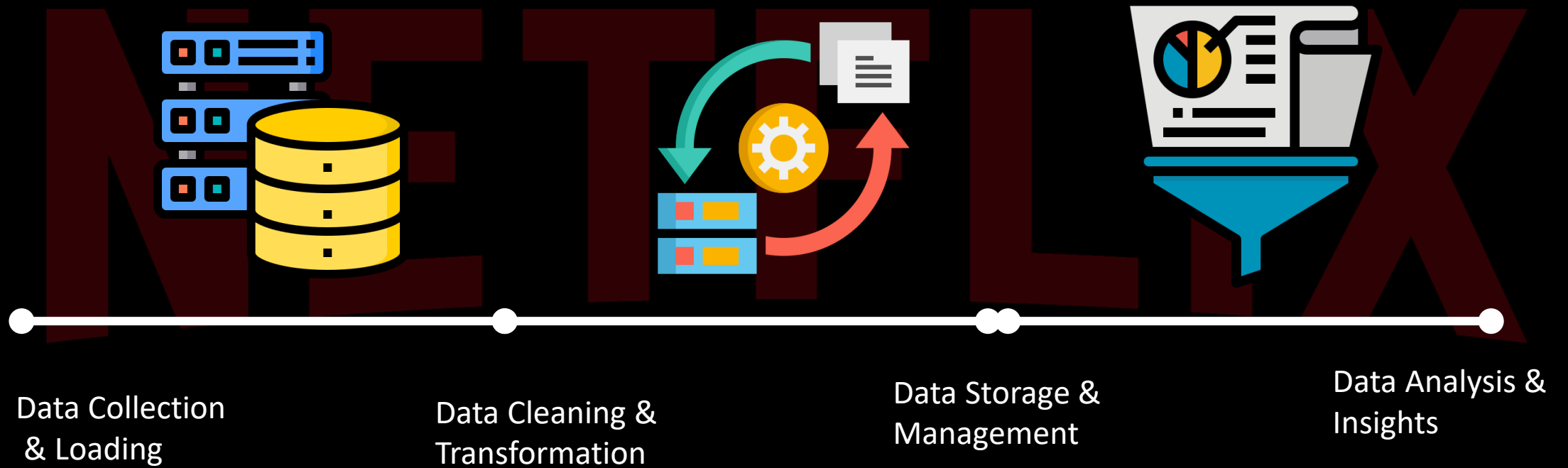
Objective

- **Analyze the distribution of content types (movies vs TV shows).**
- **Identify the most common ratings for movies and TV shows.**
- **List and analyze content based on release years, countries, and durations.**
- **Explore and categorize content based on specific criteria and keywords.**

Scope

- **Enhance Data Quality and Consistency Deepen Understanding of Viewer Preferences Optimize Operational Efficiency Drive Strategic Decision-Making.**

Workflow



Data Collection & Loading

Data Sources

- Netflix dataset containing information about movies, TV shows, directors, cast, genres, countries, duration, and ratings.

Data Loading Process

- Data was initially loaded into Python for preprocessing and then imported into the raw data layer of SQL Server.

Data Cleaning

Enhancements in Netflix Data Table

Changed Data Type of Title Column

- Converted 'title' column from 'varchar' to 'nvarchar' to support foreign characters.
- Ensured visibility and correct representation of all characters in titles.

Adjusted Column Lengths

- Determined optimal column lengths using Python to improve loading speed.
- Changed from max length to required length for efficient data handling.

Set Primary Key for show _i d

- Identified 'show _id' as unique across the dataset.
- Established 'show _id' as the primary key to ensure data integrity and uniqueness.

DATA CLEANING

```
CREATE TABLE Netflix
```

```
(
```

```
    show _id VARCHAR(max) primary key,  
    type   VARCHAR(max) NULL,  
    title   VARCHAR(max) NULL,  
    director VARCHAR(max) NULL,  
    casts   VARCHAR(max) NULL,  
    country VARCHAR(max) NULL,  
    date_added VARCHAR(max) NULL,  
    release _year INT NULL,  
    rating   VARCHAR(max) NULL,  
    duration VARCHAR(max) NULL,  
    listed _in VARCHAR(max) NULL,  
    description VARCHAR(max) NULL
```

```
);
```



```
CREATE TABLE Netflix
```

```
(
```

```
    show _id VARCHAR(10) primary key,  
    type   VARCHAR(10) NULL,  
    title   NVARCHAR(200) NULL,  
    director VARCHAR(250) NULL,  
    casts   VARCHAR(1000) NULL,  
    country VARCHAR(150) NULL,  
    date_added VARCHAR(120) NULL,  
    release _year INT NULL,  
    rating   VARCHAR(10) NULL,  
    duration VARCHAR(10) NULL,  
    listed _in VARCHAR(100) NULL,  
    description VARCHAR(500) NULL
```

```
);
```


Data Analysis

Business Problems and Solutions

- Count the number of Movies vs TV Shows

```
SELECT  
Type , count(*) as total_content  
FROM Netflix  
group by type;
```

	type	total_content
▶	Movie	55
	TV Show	45

- Find the top 5 countries with the most content on Netflix

```
SELECT country , COUNT(*) AS content_count  
FROM Netflix  
WHERE country IS NOT NULL  
GROUP BY country  
ORDER BY content_count DESC  
LIMIT 5;
```

	country	content_count
▶		41
	United States	17
	Japan	13
	India	6
	United Kingdom	5

- Find the most common rating for movies and TV shows

```
SELECT
  type,
  rating AS most _ frequent _ rating
FROM Netflix
GROUP BY type, rating
HAVING COUNT(*) = (
  SELECT MAX(rating _ count)
  FROM (
    SELECT COUNT(*) AS rating _ count
    FROM Netflix NT
    WHERE Nt. type = Netflix . type
    GROUP BY rating
  ) AS subquery
);
```

	country	content_count
▶		41
	United States	17
	Japan	13
	India	6
	United Kingdom	5

- List All Movies Released in a Specific Year (e.g., 2020)

```
SELECT *
FROM Netflix
WHERE release _ year = 2020;
```

show_id	type	title	director	cast	country	date_added	release_year	rating
s1	Movie	Dick Johnson Is Dead	Kirsten Johnson		United States	September 25, 2021	2020	PG-13
s17	Movie	Europe's Most Dangerous Man: Otto Skorzeny i...	Pedro de Echave Garc�a, Pablo Azor�n Williams			September 22, 2021	2020	TV-MA
s18	TV Show	Falsa identidad		Luis Ernesto Franco, Camila Sodi, Sergio Goyri, ...	Mexico	September 22, 2021	2020	TV-MA
s33	TV Show	Sex Education		Asa Butterfield, Gillian Anderson, Ncuti Gatwa, ...	United Kingdom	September 17, 2021	2020	TV-MA
s35	TV Show	Tayo and Little Wizards		Dami Lee, Jason Lee, Bommie Catherine Han, J...		September 17, 2021	2020	TV-Y7
s48	TV Show	The Smart Money Woman	Bunmi Ajakaiye	Osas Ighodaro, Ini Dima-Okojie, Kemi Lala Akind...		September 16, 2021	2020	TV-MA
s79	Movie	Tughlaq Durbar	Delhiprasad Deenadayalan	Vijay Sethupathi, Parthiban, Raashi Khanna		September 11, 2021	2020	TV-14
s85	Movie	Omo Ghetto: the Saga	JJC Skillz, Funke Akindele	Funke Akindele, Ayo Makun, Chioma Chukwuka ...	Nigeria	September 10, 2021	2020	TV-MA

- Identify the Longest Movie

```
select * from Netflix
where type='movie' and duration=(select max(duration) from Netflix);
```

show_id	type	title	director	cast	country	date_added	release_year
s52	Movie	InuYasha the Movie 2: The Castle Beyond the L...	Toshiya Shinohara	Kappei Yamaguchi, Satsuki Yukino, Mieko Harad...	Japan	September 15, 2021	2002
s53	Movie	InuYasha the Movie 3: Swords of an Honorable ...	Toshiya Shinohara	Kappei Yamaguchi, Satsuki Yukino, Koji Tsujitani...	Japan	September 15, 2021	2003

- Find Content Added in the Last 5 Year

```
SELECT*
FROM Netflix
WHERE STR_TO_DATE(date _ added, '%M %d, %Y') >= CURDATE() - INTERVAL 5 YEAR;
```

show_id	type	title	director	cast	country	date_added	release_year
s1	Movie	Dick Johnson Is Dead	Kirsten Johnson		United States	September 25, 2021	2020
s2	TV Show	Blood & Water		Ama Qamata, Khosi Ngema, Gail Mablane, Tha...	South Africa	September 24, 2021	2021
s3	TV Show	Ganglands	Julien Lederqc	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...		September 24, 2021	2021
s4	TV Show	Jailbirds New Orleans				September 24, 2021	2021
s5	TV Show	Kota Factory		Mayur More, Jitendra Kumar, Ranjan Raj, Alam ...	India	September 24, 2021	2021
s6	TV Show	Midnight Mass	Mike Flanagan	Kate Siegel, Zach Gilford, Hamish Linklater, Hen...		September 24, 2021	2021

- List all TV shows with more than 5 seasons

```
SELECT title, duration from Netflix WHERE type = 'TV Show' AND duration > 5
order by duration desc;
```

	title	duration
►	The Great British Baking Show	9 Seasons
	Saved by the Bell	9 Seasons
	Nailed It	6 Seasons
	Numberblocks	6 Seasons
	Lucifer	6 Seasons

- Count the Number of Content Items in Each Genre

```
SELECT listed _ in, COUNT(*) AS content _ count
FROM Netflix
GROUP BY listed _ in;
```

listed_in	content_count
Documentaries	2
International TV Shows, TV Dramas, TV Mysteries	2
Crime TV Shows, International TV Shows, TV Ac...	2
Docuseries, Reality TV	1
International TV Shows, Romantic TV Shows, TV...	2
TV Dramas, TV Horror, TV Mysteries	1
Children & Family Movies	6

- List All Movies that are Documentaries

```
SELECT * FROM Netflix
WHERE listed _ in LIKE '%Documentaries';
```

show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in
s1	Movie	Dick Johnson Is Dead	Kirsten Johnson		United States	September 25, 2021	2020	PG-13	90 min	Documentaries
s46	Movie	My Heroes Were Cowboys	Tyler Greco			September 16, 2021	2021	PG	23 min	Documentaries

Insights & Results

Key Findings

- **Content Preferences:** Highlighted the popularity of comedy movies in the United States, guiding localized content acquisition strategies.
- **Directorial Excellence:** Recognized prolific directors and their impact on Netflix's content diversity and viewer engagement.
- **Genre Analysis:** Provided insights into viewer preferences based on genre and average duration, informing content curation strategies.

Business Impact

Enhanced Content Relevance

- Targeting popular genres like comedy in the United States improves viewer engagement.

Strategic Partnerships

- Collaborating with top directors diversifies content and attracts broader audiences.

Operational Efficiency

- Optimizing content length enhances viewer satisfaction and retention



Thank you!

For exploring this presentation on the Netflix data analysis project. Your interest and engagement are greatly appreciated as we continue to uncover valuable insights from the data.