

Data Analyst
“Netflix Movie and TV Show”



Dibuat Oleh:

Kelompok 19

11423058 | Rakhel Franiska Tampubolon

11423060 | Gracia Anggraini Manik

11423064 | Yenny Angelita Gurning

INSTITUT TEKNOLOGI DEL

MATERI PRAKTIKUM

Semester V Tahun Ajar 2025/2026

Daftar Isi

Daftar Isi	1
Daftar Gambar	2
BAB 1 Pendahuluan.....	3
1.1 Latar Belakang	3
1.2 Tujuan.....	3
Bab 2 Tahapan CRISP-DM	4
2.1 Business Understanding	5
2.2 Data Understanding.....	6
2.3 Data Preparation.....	7
2.4 Modelling	12
2.5 Evaluation.....	23
2.6 Deployment	26
Bab 3 Hasil eksplorasi dan analisis.....	28
3.1 Komposisi Konten: Film vs TV Show	28
3.2 Distribusi Konten Berdasarkan Negara.....	30
3.3 Distribusi Genre Terpopuler.....	31
3.4 Tren Penambahan Konten per Tahun.....	32
3.5 Kesimpulan Hasil Analisis.....	33
Daftar Pustaka.....	34

Daftar Gambar

Gambar 2. 1 Tahapan CRISP-DM	4
Gambar 2. 2 Director (Clean)	8
Gambar 2. 3 Cast (Clean).....	8
Gambar 2. 4 Country (Clean).....	9
Gambar 2. 5 Nilai Null Menjadi Unknown	9
Gambar 2. 6 Mengubah tipe data kolom Release Year di Tableau	10
Gambar 2. 7 Movie Duration (Minutes) Calculation Field.....	11
Gambar 2. 8 TV Show SeasonsCalculation Field.....	11
Gambar 2. 9 Visualisasi Pertumbuhan Film vs Serial TV di Netflix dari Waktu ke Waktu	13
Gambar 2. 10 Visualisasi Forecast Pertumbuhan Produksi Konten Netflix (2021-2023).....	14
Gambar 2. 11 Visualisasi Rata-Rata Tahun Release	15
Gambar 2. 12 Visualisasi Distribusi Konten Berdasarkan Negara	16
Gambar 2. 13 Visualisasi Distribusi Jumlah Film dan Serial TV di Netflix	17
Gambar 2. 14 Visualisasi Analisis Rating	18
Gambar 2. 15 Visualisasi Top 10 Genre	19
Gambar 2. 16 Visualisasi 10 Negara Dengan Jumlah Film Dan Acara TV	20
Gambar 2. 17 Visualisasi Forecast Untuk Pertumbuhan Jumlah Tayangan Netflix Tahun 2021-2023.....	21
Gambar 2. 18 Visualisasi Judul Film/TV Shows Terakhir Ditambahkan Pada Netflix	22
Gambar 2. 19 Visualisasi Total Fim & TV Show.....	22
Gambar 2. 20 Dashboard	27
Gambar 3. 1 Perbandingan jumlah Film dan TV Show di Netflix	29
Gambar 3. 2 Total Movies & TV Shows berdasarkan Negara	30
Gambar 3. 3 Top 10 Genre Terpopuler di Netflix	31
Gambar 3. 4 Tren Penambahan Film dan TV Show di Netflix	32

BAB 1

Pendahuluan

1.1 Latar Belakang

Netflix merupakan salah satu *platform* streaming media dan video paling populer. *Platform* netflix memiliki film dan acara TV lebih dari 8000 dan pada pertengahan tahun 2021, *platform* ini memiliki lebih dari 200 juta pelanggan dari seluruh dunia. Dataset tabular dari *Netflix Movie and TV Show* berisi daftar semua film dan dan acara TV yang tersedia pada platform, beserta detail dari film dan acara TV seperti judul (*title*), pemeran (*cast*), sutradara(*director*), negara asal film atau acara TV(*country*), tanggal penambahan film atau acara TV ke *platform* (*date_added*), tahun rilis film(*release_year*), *rating*, durasi film atau acara TV(*duration*), kategori film (*listed_in*) dan deskripsi dari film atau acara TV(*description*).

Penelitian oleh *Journal of Engineering Research and Reports (2023)* menunjukkan bahwa analisis data konten Netflix dapat memberikan insight penting terkait tren genre, negara produksi, dan pola perilaku pengguna melalui pendekatan eksplorasi data dan visualisasi.[1]

Melalui analisis dataset *Netflix Movie and TV Show*, maka dapat menghasilkan berbagai informasi yang berguna seperti negara mana yang paling banyak memproduksi atau memiliki konten pada *platform Netflix*, jenis film atau acara TV yang mirip, aktor yang bekerja dengan sutradara, perbandingan pertumbuhan film dengan acara TV sepanjang waktu dan kategori yang semakin meningkat ataupun yang menurun ketenarannya di tiap tahun.

1.2 Tujuan

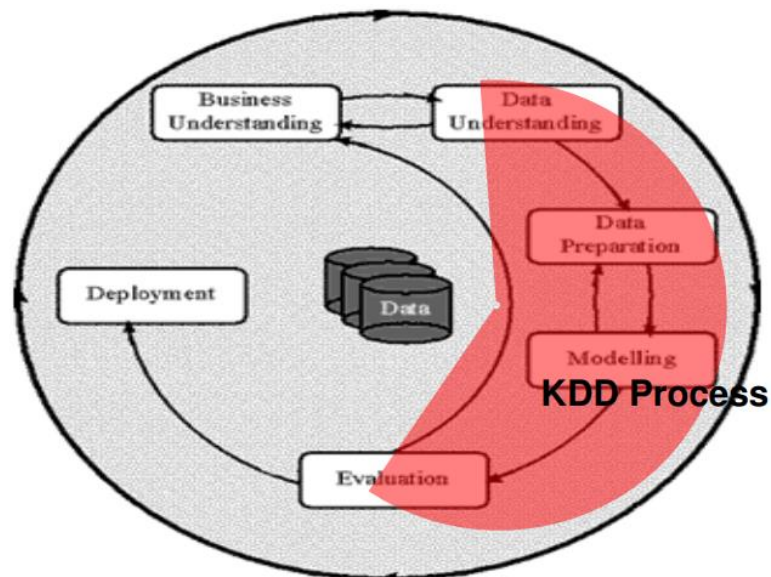
Tujuan untuk melakukan analisis ini adalah untuk memahami isi dan pola dataset *Netflix Movie and TV Show*. Pada proyek ini analisis ini secara khusus membahas tentang:

1. Memahami konten apa saja yang tersedia di berbagai negara.
2. Mengidentifikasi konten yang serupa, dengan menyamakan fitur berbasis *text*.
3. Menganalisis jaringan aktor/sutradara.
4. Menganalisis apakah *platform Netflix* lebih fokus pada film atau acara TV dalam beberapa tahun terakhir.

Bab 2

Tahapan CRISP-DM

Metodologi CRISP-DM (Cross Industry Standard Process for Data Mining) merupakan kerangka kerja yang banyak digunakan dalam proyek data mining dan analitik karena sifatnya yang sistematis, terstruktur, dan fleksibel [2]. Metodologi ini membagi keseluruhan proses menjadi enam fase utama: Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, dan Deployment. Setiap fase saling berkaitan dan memungkinkan iterasi, sehingga memudahkan penyesuaian dan perbaikan di sepanjang perjalanan proyek. Tahapan CRISP-DM dapat dilihat pada Gambar 2.1 berikut.



Gambar 2. 1 Tahapan CRISP-DM

sumber:<https://didawiki.cli.di.unipi.it/lib/exe/fetch.php/dm/crisp-dm.pdf>

Pada bab ini, setiap fase akan dijelaskan secara ringkas agar pembaca dapat memahami alur proses serta peranan penting masing-masing fase dalam proyek analitik data. Fase Business Understanding akan menjadi landasan bagi analisis lebih lanjut, karena di fase inilah masalah bisnis dirumuskan dan diubah menjadi permasalahan analitik yang konkret. Pemahaman yang mendalam tentang konteks bisnis sangat penting, karena hal ini akan mempengaruhi seluruh proses pengolahan data dan pengambilan keputusan yang akan dilakukan.

2.1 Business Understanding

Tahapan Business Understanding merupakan langkah awal dalam metodologi CRISP-DM (Cross Industry Standard Process for Data Mining) yang berfokus pada pemahaman konteks bisnis dan penentuan permasalahan analitik yang ingin diselesaikan menggunakan data. Tahap ini penting agar analisis yang dilakukan relevan dengan tujuan bisnis dan menghasilkan insight yang bermakna.

Pada proyek ini digunakan dataset “Netflix Movies and TV Shows” yang berisi informasi terkait berbagai film dan serial televisi di platform Netflix, mencakup atribut seperti jenis konten, negara asal, tahun rilis, durasi, genre, dan rating. Dataset ini memungkinkan dilakukannya analisis deskriptif untuk memahami tren dan karakteristik konten yang tersedia di Netflix.

Permasalahan analitik bisnis yang ingin dijawab melalui analisis ini meliputi:

1. Bagaimana perkembangan jumlah film dan serial TV di Netflix dari tahun ke tahun?
2. Bagaimana Prediksi pertumbuhan jumlah konten Netflix hingga tahun 2025 berdasarkan tren historis?
3. Bagaimana distribusi rata-rata tahun rilis konten di netflix?
4. Negara mana yang paling banyak memproduksi konten di Netflix?
5. Bagaimana perbandingan jumlah film dan serial TV dalam katalog Netflix secara keseluruhan?
6. Bagaimana distribusi rating dari konten yang tersedia di Netflix?
7. Negara mana saja yang menempati posisi 10 besar dalam jumlah film dan acara Tv di Netflix?
8. Bagaimana Prediksi jumlah tayangan Netflix dalam jangka pendek (2021-2023)?
9. Kapan dan seberapa sering konten baru terakhir kali ditambahkan di Netflix ?

Berdasarkan permasalahan tersebut, tujuan dari tahap ini adalah untuk memahami pola dan tren konten Netflix yang dapat mendukung pengambilan keputusan strategis, khususnya dalam aspek pengelolaan dan pengembangan konten di masa mendatang.

2.2 Data Understanding

Tahapan ini bertujuan untuk memahami isi dan karakteristik awal dari dataset Netflix Movies and TV Shows sebelum dilakukan proses pembersihan dan analisis lebih lanjut. Dataset ini berisi informasi terkait daftar film dan acara TV yang tersedia di platform Netflix, mencakup berbagai atribut seperti judul konten, jenis tayangan (Movie atau TV Show), nama sutradara, pemeran utama, negara asal produksi, tanggal penambahan ke Netflix, tahun rilis, rating penonton, durasi tayangan, genre atau kategori, serta deskripsi singkat.

Eksplorasi awal dilakukan menggunakan Tableau melalui tampilan Data Source untuk meninjau tipe data dan jumlah entri pada setiap kolom. Berdasarkan pengamatan, ditemukan bahwa beberapa kolom seperti director, cast, dan country memiliki nilai kosong (*missing value*) menandakan bahwa tidak semua data tersedia secara lengkap.

Director	Cast	Country
<i>null</i>	<i>null</i>	<i>null</i>
<i>null</i>	<i>null</i>	<i>null</i>
<i>null</i>	<i>null</i>	<i>null</i>
<i>null</i>	<i>null</i>	<i>null</i>

Selain itu, terdapat tipe data yang tidak sesuai, seperti pada kolom 'Release Year', yang seharusnya bertipe Date & Time, namun saat ini ditetapkan sebagai Number (whole).

Number (decimal)	
✓ Number (whole)	
Date & Time	
Date	
String	
Boolean	
✓ Default	
Geographic Role	

Release Year	Rating	Dur
2020	PG-13	90
2021	TV-MA	2 S
2021	TV-MA	1 S

Temuan ini menjadi data penting dalam menentukan langkah pada tahap Data Preparation, sehingga dapat digunakan secara optimal dalam proses analisis dan visualisasi berikutnya.

2.3 Data Preparation

Data preparation (persiapan data), juga dikenal sebagai pengolahan data (data wrangling), adalah proses di mana data diubah dari representasi yang ada menjadi bentuk yang sesuai untuk analisis [3]. Tahapan Data Preparation bertujuan untuk menyiapkan data agar siap digunakan dalam proses analisis dan visualisasi. Pada tahap ini dilakukan proses pembersihan (*data cleaning*), transformasi (*data transformation*), serta integrasi data (*data integration*) untuk memastikan bahwa data yang digunakan sudah akurat, konsisten, dan sesuai dengan kebutuhan analisis.

Proses persiapan data dilakukan pada dataset “**Netflix Movies and TV Shows**” dengan langkah-langkah sebagai berikut:

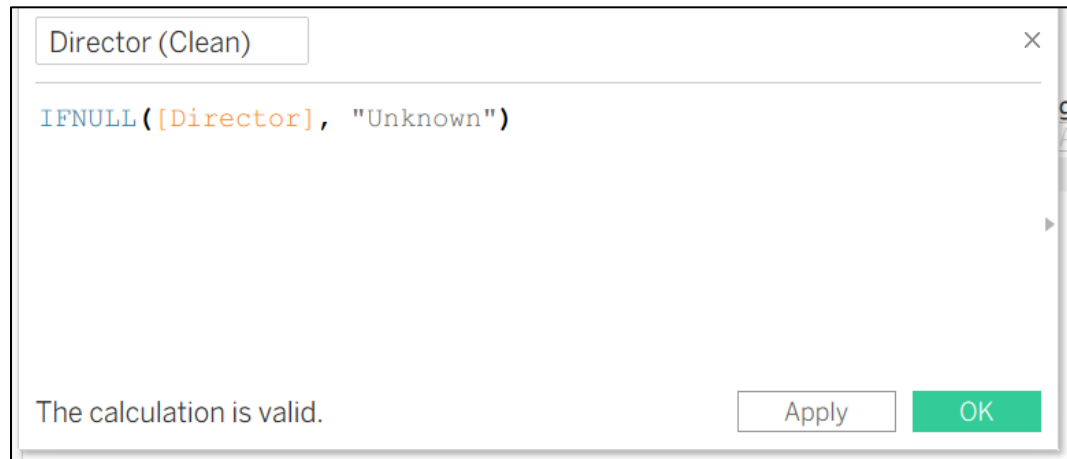
1. Handling Missing Values

Pada tahap data handling, dilakukan pemeriksaan data duplikat untuk memastikan tidak ada entri yang sama muncul lebih dari satu kali. Pemeriksaan dilakukan pada kolom `show_id` sebagai identifier unik untuk setiap konten.

Hasil pemeriksaan menunjukkan bahwa tidak terdapat data duplikat, sehingga seluruh data dapat digunakan untuk proses analisis.

Selain itu, nilai kosong (missing value) pada kolom `director`, `cast`, dan `country` diganti dengan teks “Unknown” menggunakan fungsi `IFNULL()` di Tableau untuk menjaga konsistensi dataset. Berikut langkah yang dilakukan untuk menangani hal tersebut.

- a. Nilai kosong pada kolom `director` diganti dengan teks “Unknown” untuk menjaga konsistensi data tanpa menghapus baris yang valid. Dapat dilihat pada Gambar 2.2 berikut.



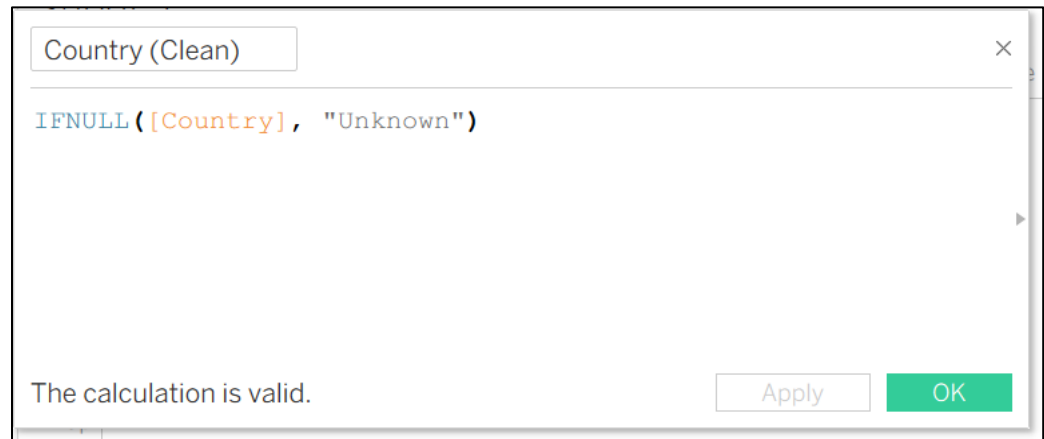
Gambar 2. 2 Director (Clean)

- b. Nilai kosong pada kolom cast diganti dengan teks “Unknown” untuk menjaga konsistensi data tanpa menghapus baris yang valid. Dapat dilihat pada Gambar 2.3 berikut.



Gambar 2. 3 Cast (Clean)

- c. Nilai kosong pada kolom country diisi dengan “Unknown” agar tetap bisa dianalisis dalam kategori tersendiri. Dapat dilihat pada Gambar 2.4 berikut.



Gambar 2. 4 Country (Clean)

- d. Kolom dengan nilai kosong terlalu banyak tetap dipertahankan karena masih relevan secara analitis. Sehingga terdapat 3 kolom baru yang dimana nilai null sudah diubah menjadi Unknown. Dapat dilihat pada Gambar 2.5 berikut.

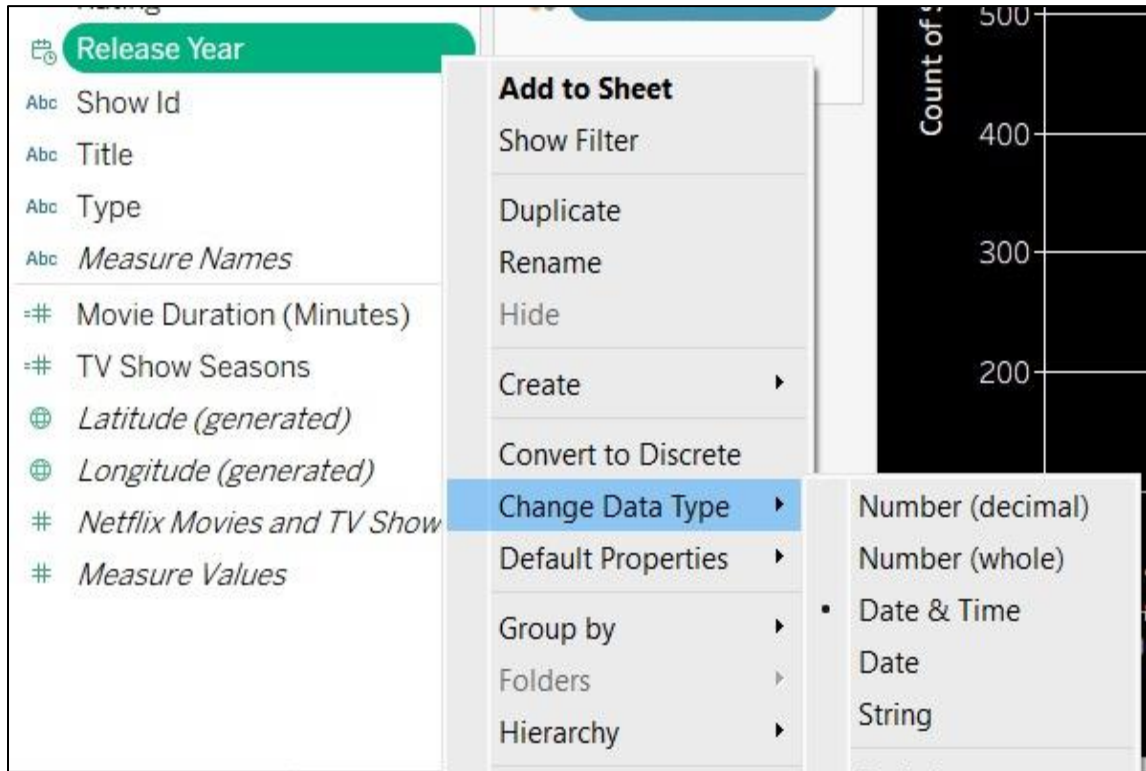
Director (Clean)	Cast (Clean)	Country (Clean)
Unknown	Ama Qamata, Khosi Ngema, ...	South Africa
Unknown	Unknown	Unknown
Unknown	Mayur More, Jitendra Kumar, ...	India

Gambar 2. 5 Nilai Null Menjadi Unknown

2. Data Transformation

Tahap transformasi dilakukan untuk menyesuaikan format dan struktur data agar lebih mudah dianalisis di Tableau. Proses transformasi yang dilakukan meliputi:

Mengganti tipe data pada kolom tertentu:

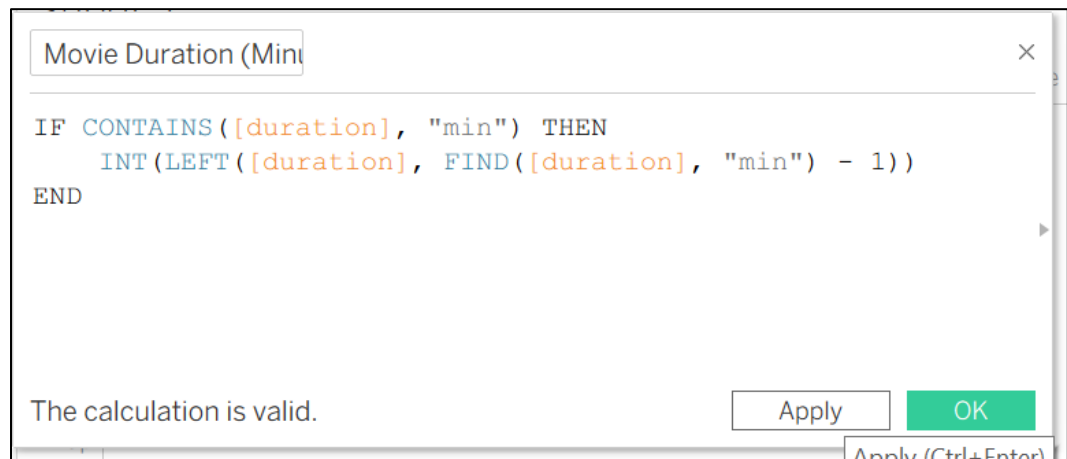


Gambar 2. 6 Mengubah tipe data kolom Release Year di Tableau

Proses perubahan tipe data dilakukan pada kolom Release Year, yang semula bertipe Number (Whole) diubah menjadi Date & Time. Langkah ini dilakukan agar Tableau dapat mengenali kolom tersebut sebagai data waktu, sehingga dapat digunakan untuk analisis berbasis tahun secara lebih akurat dan mendukung visualisasi temporal (seperti tren atau perkembangan dari waktu ke waktu).

Memisahkan nilai pada kolom duration menjadi dua bentuk baru untuk membedakan jenis konten:

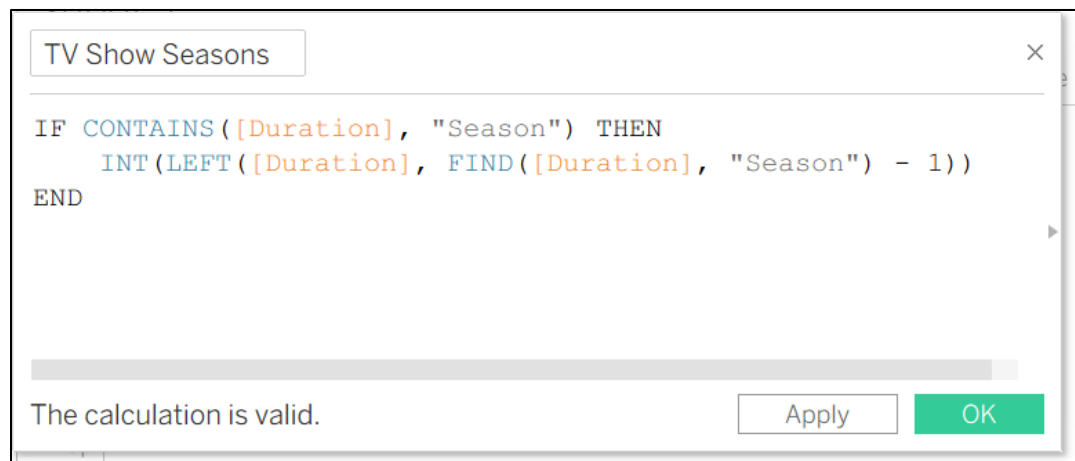
- a. Movie Duration (Minutes) untuk konten film, menggunakan formula:



Gambar 2. 7 Movie Duration (Minutes) Calculation Field

Fungsi ini memeriksa apakah string yang terdapat dalam kolom duration mengandung kata "min", yang menunjukkan bahwa durasi yang dicatat menggunakan satuan menit.

- b. TV Show Seasons untuk konten serial, menggunakan formula:



Gambar 2. 8 TV Show SeasonsCalculation Field

Fungsi ini mengambil substring dari kolom duration mulai dari kiri hingga karakter sebelum kata "min", sehingga menghasilkan angka yang merepresentasikan durasi dalam menit.

Dengan dua field ini, analisis durasi dapat dilakukan secara terpisah antara film dan serial. Transformasi ini membuat data lebih terstruktur, mudah dibaca, serta siap digunakan dalam berbagai jenis visualisasi di Tableau.

3. Data Integration

Karena seluruh data sudah terdapat dalam satu file CSV, tidak diperlukan penggabungan antar sumber data. Namun, pada tahap ini dilakukan penyelarasan antar kolom agar data bisa diolah dengan baik di Tableau, seperti:

- a. Menyamakan format teks (misalnya huruf kapital diubah menjadi format *Title Case*).
- b. Menghapus spasi berlebih dan karakter khusus.
- c. Memastikan semua kolom memiliki nama yang jelas dan konsisten.

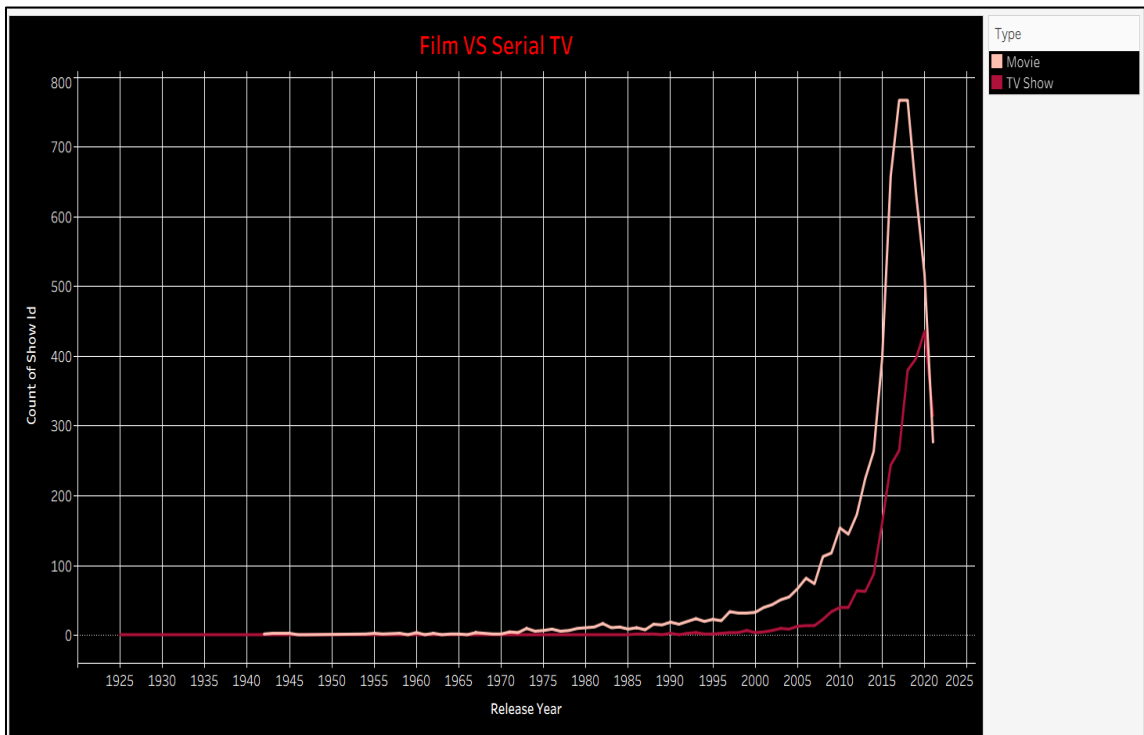
4. Hasil Akhir Persiapan Data

Setelah seluruh proses data handling, data transformation, dan data integration dilakukan di Tableau, dataset kini telah bersih dari nilai kosong yang mengganggu, memastikan bahwa analisis dapat dilakukan dengan lebih akurat. Selain itu, dataset memiliki format tipe data yang sesuai, yang memudahkan dalam pemrosesan lebih lanjut. Dalam proses ini, juga ditambahkan kolom-kolom baru seperti Year Added, Movie Duration (Minutes), dan TV Show Seasons, yang memberikan informasi tambahan yang berguna untuk analisis. Dengan semua perbaikan dan penambahan ini, dataset kini siap digunakan untuk proses Descriptive Analytics dan pembuatan dashboard interaktif pada tahap berikutnya.

2.4 Modelling

Bagian ini akan menampilkan visualisasi data yang bertujuan untuk memberikan gambaran menyeluruh tentang mengenai pola dan tren yang terdapat pada dataset. Pada tiap grafik dibuat untuk menyoroti beberapa bagian berbeda dari data seperti distribusi, perbandingan dan hubungan antar variabel.

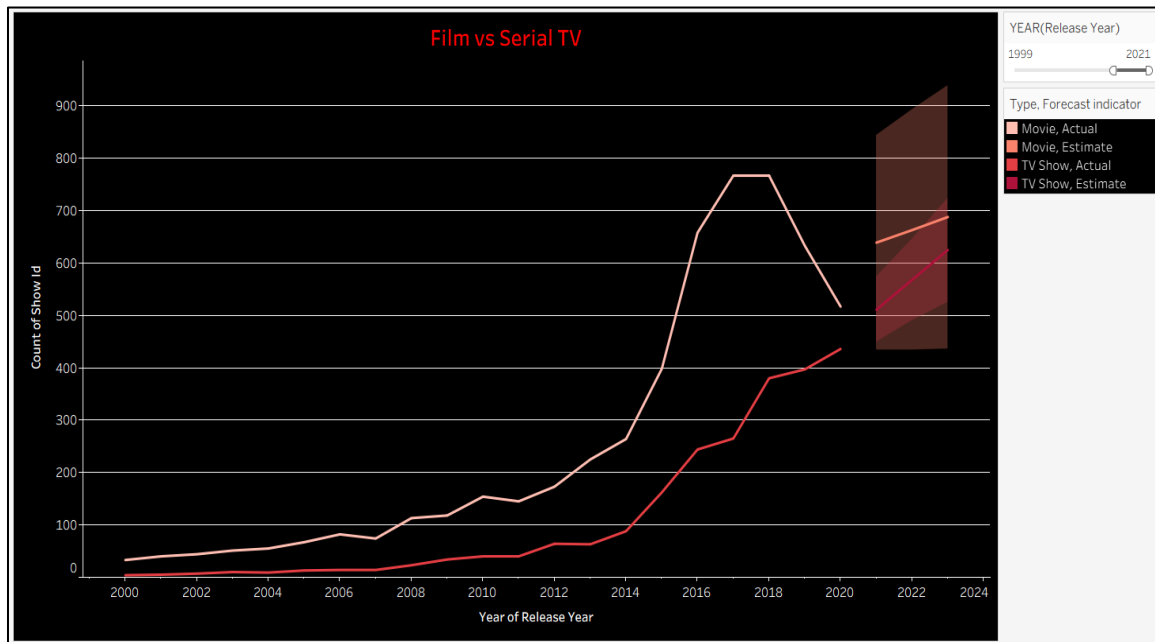
a. Pertumbuhan Film vs Serial TV di Netflix dari Waktu ke Waktu



Gambar 2. 9 Visualisasi Pertumbuhan Film vs Serial TV di Netflix dari Waktu ke Waktu

Visualisasi pertumbuhan film dan serial TV menunjukkan peningkatan pesat setelah tahun 2000, dengan puncak pertumbuhan terjadi sekitar tahun 2017-2019. Film menunjukkan jumlah yang konsisten lebih tinggi dibandingkan acara TV di setiap periode. Meskipun demikian tren kenaikan pada acara TV juga signifikan menunjukkan perluasan netflix dalam memperbanyak konten serial untuk menarik penonton yang menyukai format berkelanjutan. Setelah 2020, terjadi sedikit penurunan pada keduanya, kemungkinan akibat dampak pandemi terhadap proses produksi.

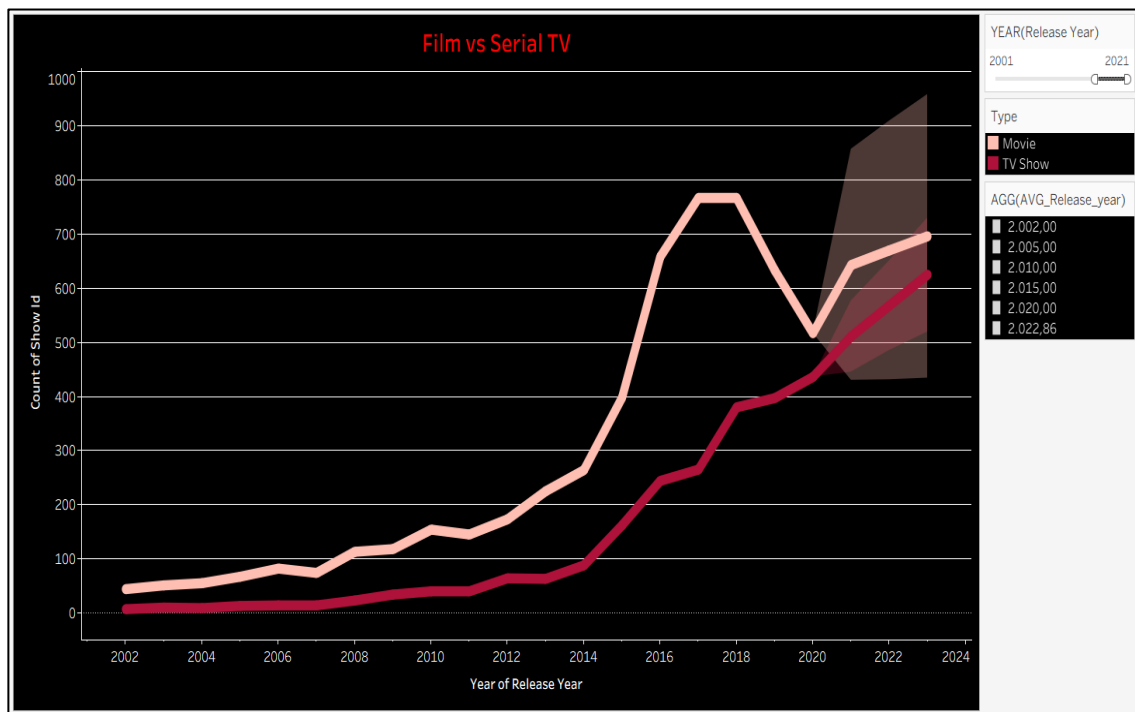
Forecast Pertumbuhan Produksi Konten Netflix (2021-2023)



Gambar 2. 10 Visualisasi Forecast Pertumbuhan Produksi Konten Netflix (2021-2023)

Forecast diatas menampilkan perbandingan tayangan serial TV dari tahun 1999 hingga 2021, serta ramalan untuk 3 tahun kedepan sampai 2024. Jumlah tayangan film meningkat secara signifikan tiap tahunnya, begitu juga dengan jumlah tayangan TV show meningkat namun laju pertumbuhannya lebih lambat dibandingkan film. Untuk film, diperkirakan jumlah tayangan akan meningkat dari sekitar 517 hingga 688 yaitu dari tahun 2020-2023, untuk serial TV diperkirakan jumlah tayangan akan meningkat dari 426 sampai 625 di tahun 2023.

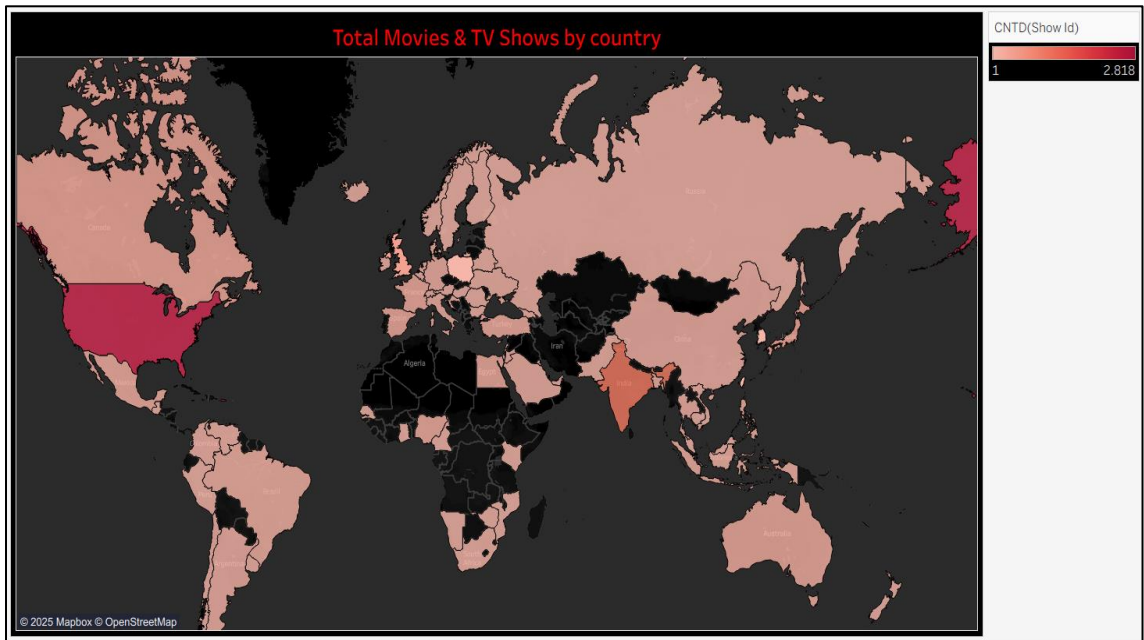
Visualisasi Rata-Rata Tahun Release



Gambar 2. 11 Visualisasi Rata-Rata Tahun Release

Pada visualisasi sebelumnya telah dijelaskan perbandingan antara film dan serial TV beserta dengan forecast nya untuk 3 tahun kedepannya. Pada bagian ini juga dijelaskan visualisasi dengan rata-rata tahun release yang dimana ketika filter diterapkan contohnya dari tahun 2001 hingga 2021 maka akan memunculkan rata-rata tahun mana saja yang paling banyak rilis.

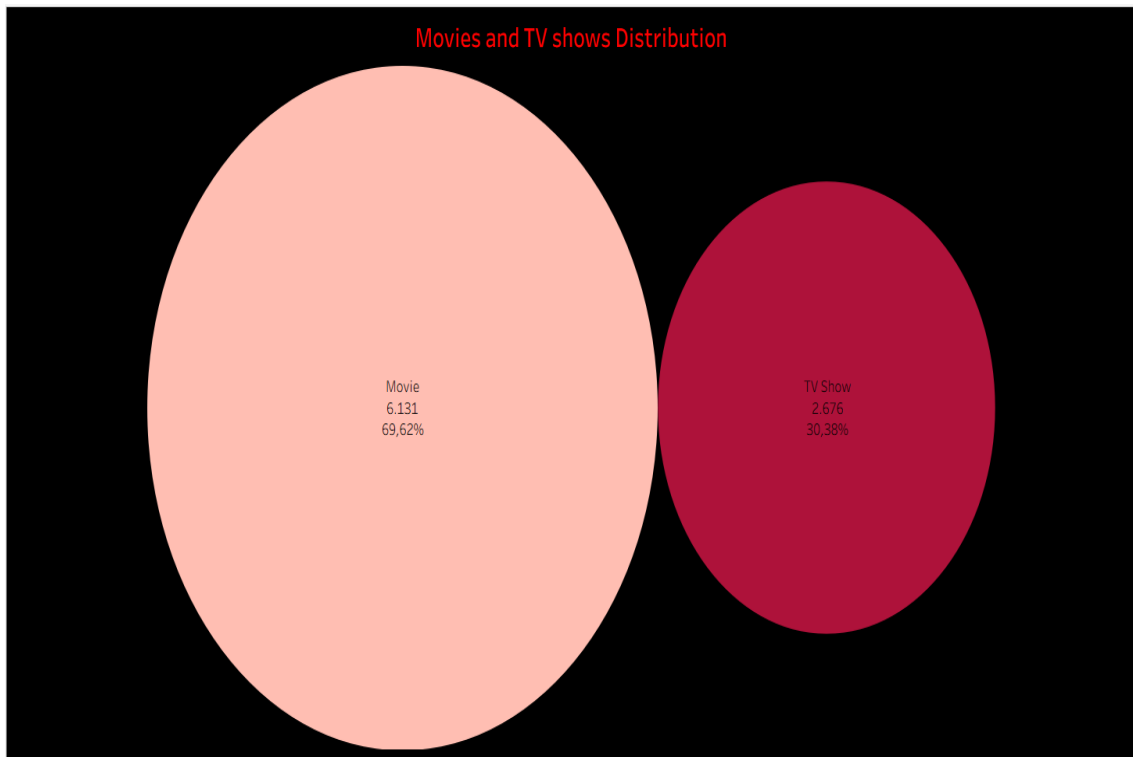
b. Distribusi Konten Berdasarkan Negara



Gambar 2. 12 Visualisasi Distribusi Konten Berdasarkan Negara

Visualisasi peta "Total Movies & TV Shows by country" berhasil mengidentifikasi negara dengan produksi konten terbanyak, yang merupakan salah satu tujuan utama analisis. Dari peta tersebut, terlihat jelas bahwa Amerika Serikat adalah produsen konten terbesar di platform Netflix.

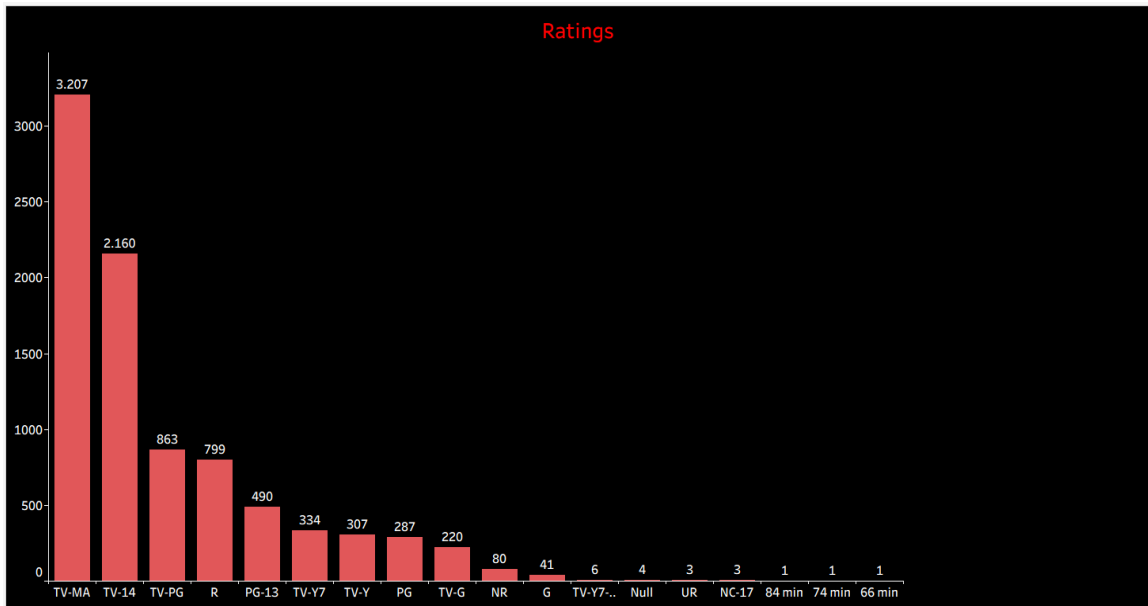
c. Distribusi Jumlah Film dan Serial TV di Netflix



Gambar 2. 13 Visualisasi Distribusi Jumlah Film dan Serial TV di Netflix

Visualisasi "Film VS Serial TV" dan "Movies and TV shows Distribution" secara langsung menjawab pertanyaan tentang fokus konten Netflix. Ditemukan bahwa Netflix lebih banyak memiliki film (69,62%) dibandingkan acara TV (30,38%). Grafik pertumbuhan sepanjang waktu juga menunjukkan bahwa jumlah film yang ditambahkan secara konsisten lebih tinggi daripada acara TV.

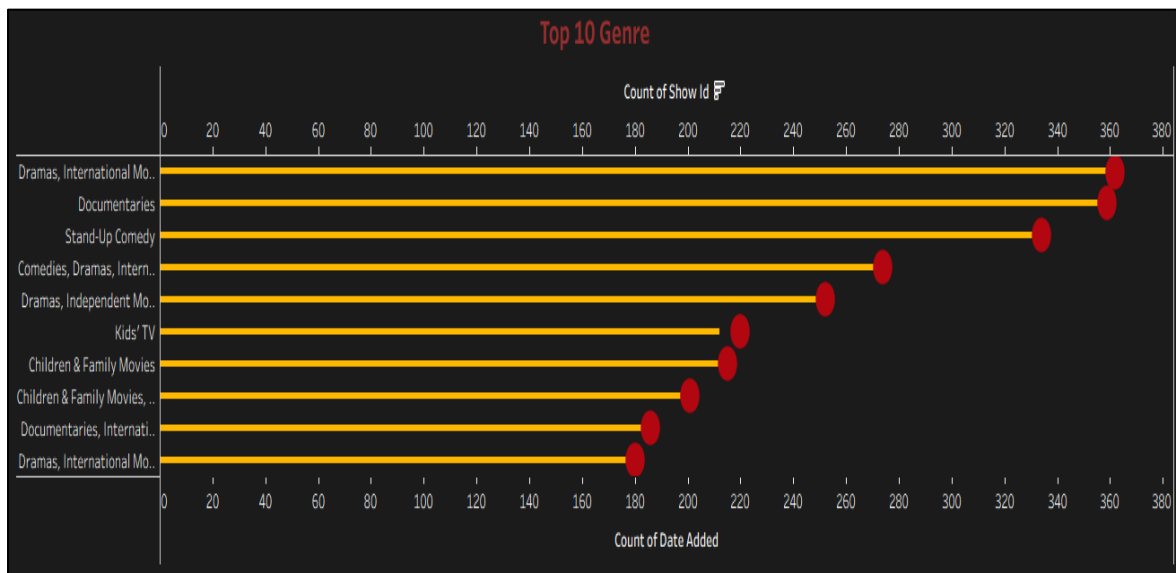
d. Analisis Rating



Gambar 2. 14 Visualisasi Analisis Rating

Grafik "Ratings" memberikan gambaran umum tentang demografi penonton yang menjadi target Netflix. Mayoritas konten memiliki rating untuk penonton dewasa (TV-MA) dan remaja (TV-14), yang memberikan pemahaman tentang jenis konten yang mendominasi platform.

e. Top 10 Genre



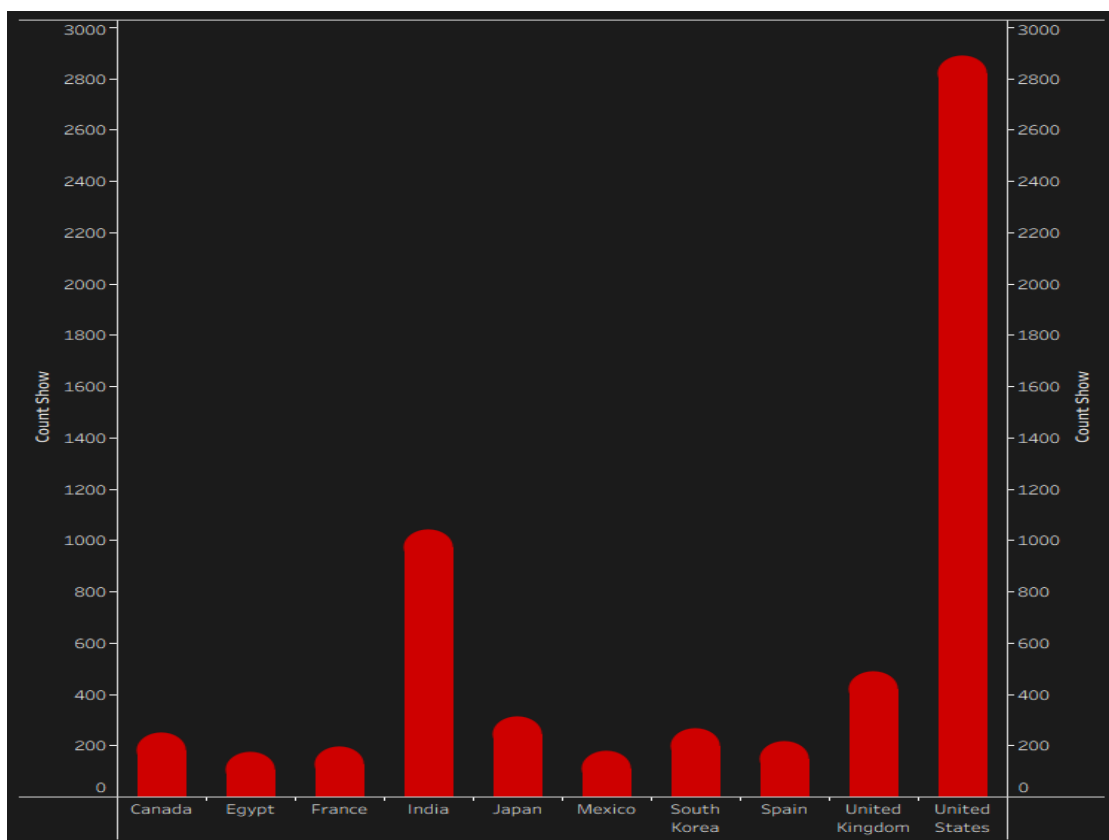
Gambar 2. 15 Visualisasi Top 10 Genre

Pada visualisasi gambar Top 10 Genre Teratas Jumlah Tayangan Terbanyak di Netflix, visualisasi ini membantu memahami genre apa yang paling dominan di platform tersebut. Dari visualisasi, terlihat bahwa genre Drama, International Movies, dan Documentaries

merupakan kategori dengan jumlah penayangan terbanyak. Pola yang ditunjukkan dalam gambar mengindikasikan bahwa Netflix banyak memproduksi genre dokumenter dan drama dibandingkan dengan genre lainnya.

Visualisasi ini menerapkan Descriptive Analytics untuk analisis statistik sederhana, khususnya pada bagian mode, yang menunjukkan nilai yang paling sering muncul dalam data tayangan. Dengan memperhatikan jumlah total tayangan dan jumlah penambahan, kita dapat melihat aktivitas penambahan konten baru dalam setiap genre. Jika titik merah mendekati ujung batang, ini menunjukkan bahwa genre tersebut sering mendapatkan penambahan tayangan baru secara konsisten.

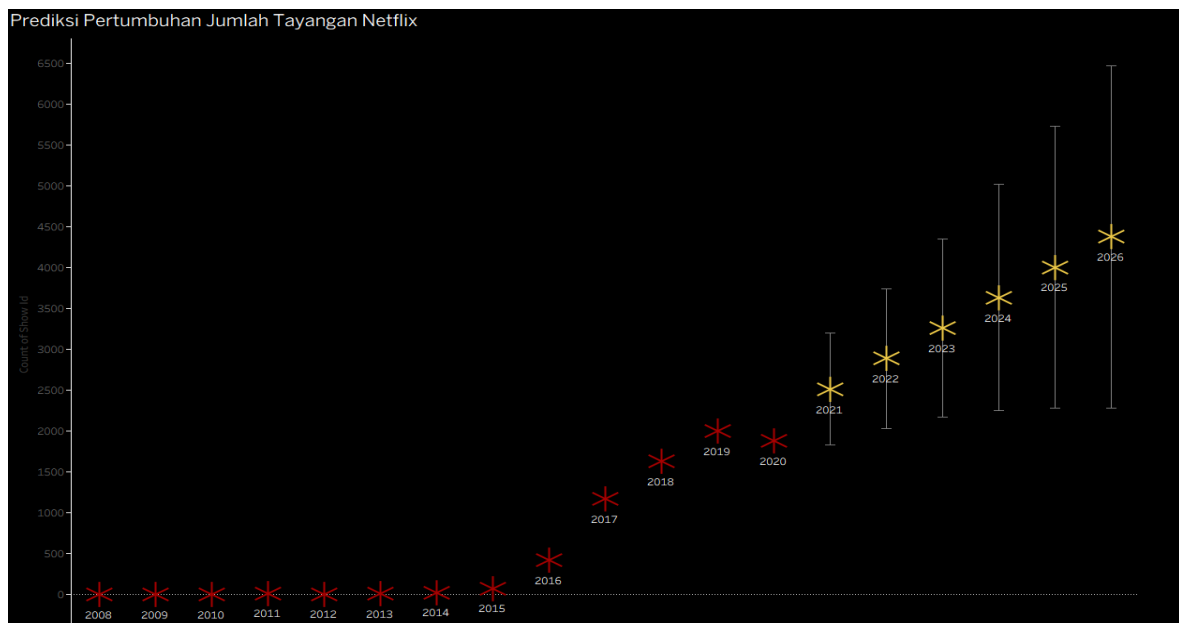
f. 10 Negara Dengan Jumlah Film Dan Acara TV



Gambar 2. 16 Visualisasi 10 Negara Dengan Jumlah Film Dan Acara TV

Pada gambar visualisasi diatas menampilkan bar chart vertikal yang menunjukkan jumlah film dan acara TV yang diproduksi oleh berbagai negara. Pada visualisasi menampilkan bahwa United States memiliki 2900 judul dimana merupakan negara dengan judul yang paling banyak dibandingkan dengan negara lain. Pada posisi kedua ada negara India, negara Kanada, Jepang dan Inggris juga memiliki jumlah judul cukup signifikan meskipun masih dibawah negara Unite State. Negara Mesir, Meksiko dan Spanyol memiliki jumlah judul yang lebih sedikit.

g. Forecast Untuk Pertumbuhan Jumlah Tayangan Netflix Tahun 2021-2023



Gambar 2. 17 Visualisasi Forecast Untuk Pertumbuhan Jumlah Tayangan Netflix Tahun 2021-2023

Pada visualisasi menampilkan pertumbuhan jumlah tayangan mulai tahun 2008 sampai 2020 dimana diberi tanda dengan warna merah, namun pada tahun 2020 terjadi penurunan jumlah tayangan Netflix. Pada warna kuning merupakan forecast atau prediksi 5 tahun kedepan terhadap jumlah pertumbuhan tayangan Netflix, pada tahun 2021 hingga 2025 menampilkan bahwa peningkatan yang bagus.

h. Judul Film/TV Shows Terakhir Ditambahkan Pada Netflix



Gambar 2. 18 Visualisasi Judul Film/TV Shows Terakhir Ditambahkan Pada Netflix

Pada gambar menampilkan visualisasi gambar film/tv shows yang ditambahkan 2 hari terakhir yaitu pada tanggal 24 september 2021 ada 9 judul yang ditambahkan dan pada tanggal 25 september hanya ada 1 judul yang ditambahkan.

i. Total Fim & TV Show



Gambar 2. 19 Visualisasi Total Fim & TV Show

Pada visualisasi menampilkan jumlah total film dan tv show yang tersedia pada Netflix. Keseluruhan data yang ditampilkan adalah hasil perhitungan dari seluruh jumlah konten yang ada pada dataset yaitu sejak tahun 2008 hingga tahun akhir 2021.

2.5 Evaluation

Tahap evaluasi ini bertujuan meninjau hasil analisis untuk memahami tren dan peluang strategis bagi Netflix. Setiap insight yang diperoleh, seperti pertumbuhan konten, distribusi negara, rating, dan genre populer, dievaluasi untuk melihat arah perkembangan industri streaming serta potensi strategi yang dapat diterapkan ke depan.

a. Pertumbuhan Film dan Serial TV di Netflix dari Waktu ke Waktu

Insight

Produksi film meningkat tajam hingga 2017, kemudian menurun setelahnya. Sebaliknya, jumlah serial TV terus meningkat stabil hingga saat ini, bahkan diperkirakan akan melampaui film di tahun-tahun mendatang. Estimasi 2022–2025 menunjukkan tren serial TV yang terus naik, sementara film tumbuh lebih lambat.

Evaluasi

- Perubahan preferensi audiens ke arah konten seriall menandakan peluang besar di segmen TV series
- Perusahaan streaming dapat mengalihkan investasi lebih banyak diproduksi serial TV karena menunjukkan tren pertumbuhan yang berkelanjutan dan potensi loyalitas penonton yang lebih tinggi
- Film tetap penting, tetapi mungkin perlu difokuskan pada kualitas daripada kuantitas

b. Distribusi Konten Berdasarkan Negara

Insight

Amerika Serikat mendominasi produksi film dan acara TV di dunia, diikuti oleh India, United Kingdom, dan Jepang. Negara-negara seperti India, Korea Selatan, dan Jepang mulai menunjukkan peningkatan signifikan dalam industri hiburan.

Evaluasi

- Memperluas jangkauan dengan meningkatkan produksi dan distribusi konten dari Asia.

- Mendorong kolaborasi internasional untuk menghasilkan konten dengan daya tarik lintas budaya.
- Mengadaptasi strategi lokalisasi agar sesuai dengan preferensi penonton di berbagai wilayah dunia.

c. Distribusi Jumlah Film dan Serial TV di Netflix

Insight

Film mendominasi dengan **69,62%** dari total konten, sementara serial TV hanya **30,38%**. Ini menunjukkan bahwa meskipun film lebih banyak, tren penonton sekarang lebih sering beralih ke serial TV (berdasarkan grafik tren sebelumnya).

Evaluasi

- Komposisi konten masih berta di film, padahal potensi pertumbuhan ada di serial.
- Bisnis sebaiknya menyeimbangkan portofolio konten, menambah lebih banyak serial original untuk mengikuti tren permintaan pasar

d. Analisis Rating

Insight

Kategori TV-MA(Konten dewasa) dan TV-14 mendominasi secara signifikan. Artinya sebagian besar konten yang tersedia diperuntukkan bagi remaja dan dewasa bukan anak-anak.

Evaluasi

- Platform dapat memperluas target pada anak dan keluarga dengan menambah PG, G dan TV-Y.
- Namun, dominasi rating dewasa menunjukkan peluang besar di konten mature dan edgy, yang biasanya lebih menarik bagi pelanggan muda dan dewasa

e. Top 10 Genre

Insight

Genre paling populer: Drama dan Documentary, diikuti oleh Stand-Up Comedy dan Comedies. Genre dengan banyak variasi “International Movies” menunjukkan minat tinggi terhadap konten lintas negara

Evaluasi

- Produksi drama internasional menjadi strategi paling relevan karena genre ini mendominasi dari sisi jumlah maupun engagement
- Dokumenter juga menjadi peluang bagus untuk meningkatkan reputasi dan diferensiasi konten edukasi di platform.

f. 10 Negara dengan Jumlah film dan Acara TV

Insight

Amerika Serikat mendominasi jauh dengan sekitar 3.000 judul, disusul oleh India, United Kingdom, dan Jepang. Negara Asia India, Korea Selatan, dan Jepang memiliki kontribusi cukup tinggi.

Evaluasi

- Dominasi AS menunjukkan masih kuatnya pengaruh Hollywood, tetapi tren globalisasi konten Asia (India dan Korea) mulai dilihat.
- Bisnis dapat memperluas jangkauan dengan memproduksi atau menayangkan lebih banyak konten lokal dari Asia, karena permintaannya meningkat pesat di pasar internasional.

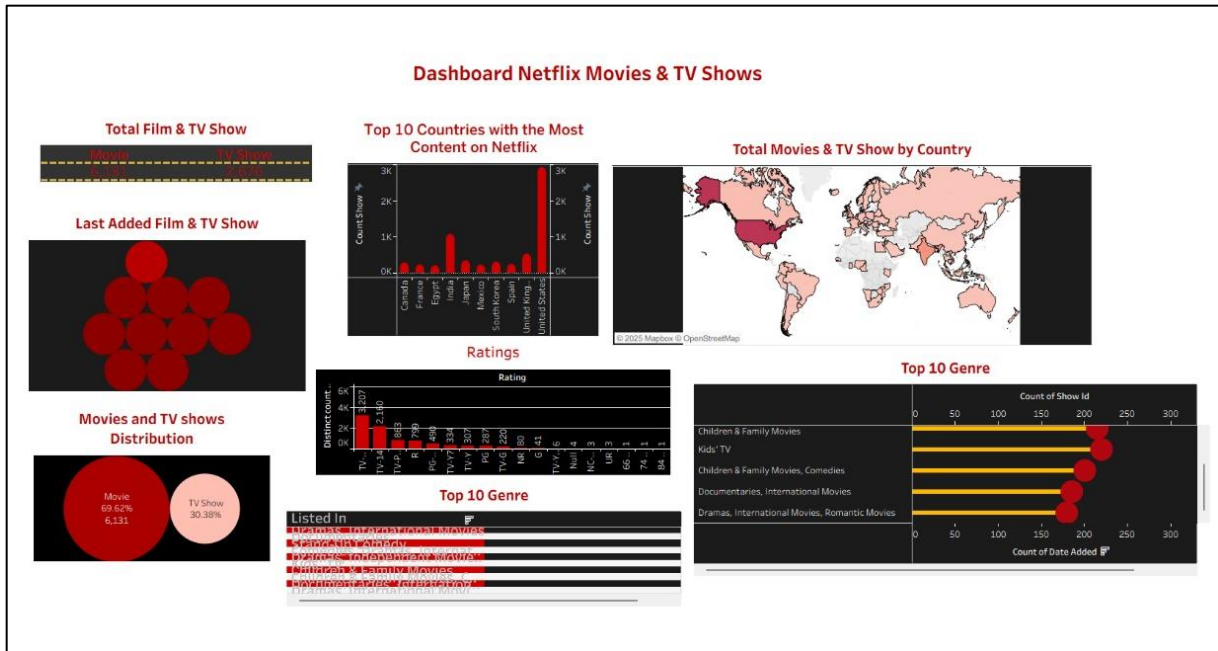
2.6 Deployment

Tahap Deployment merupakan fase akhir dalam metodologi CRISP-DM, di mana pengetahuan atau insight yang diperoleh dari hasil analisis disusun dan disajikan agar dapat digunakan oleh pihak yang berkepentingan. Tujuan utama dari tahap ini adalah memastikan bahwa hasil analisis data dapat memberikan manfaat nyata bagi proses pengambilan keputusan bisnis.

Pada proyek ini, hasil analisis dari dataset Netflix Movies and TV Shows diimplementasikan dalam bentuk dashboard interaktif pada Tableau, yang menampilkan visualisasi mengenai distribusi jenis konten, negara produsen, tren pertumbuhan konten dari tahun ke tahun, serta genre yang paling populer. Dashboard ini dirancang agar pengguna dapat secara langsung mengeksplorasi data dan memperoleh insight dengan lebih mudah dan intuitif.

Proses deployment ini dapat dikategorikan sebagai bentuk sederhana dari penerapan hasil analisis, karena berfokus pada penyajian informasi dalam format visual yang dapat digunakan untuk pengambilan keputusan strategis. Namun, dalam konteks organisasi yang lebih luas, tahap deployment juga dapat berkembang menjadi proses yang lebih kompleks misalnya dengan penerapan model analitik secara langsung dalam sistem rekomendasi, penyesuaian strategi pemasaran berbasis data, atau pembaruan katalog konten secara otomatis.

Dalam proyek ini, peran utama analis data adalah menyiapkan hasil analisis dan menyusunnya ke dalam format yang dapat dimanfaatkan oleh pengguna akhir. Dengan demikian, tahap deployment memastikan bahwa seluruh pengetahuan yang diperoleh dari proses analisis data tidak hanya berhenti pada hasil visualisasi, tetapi juga dapat diterapkan secara nyata untuk mendukung keputusan dan strategi bisnis di masa mendatang.



Gambar 2. 20 Dashboard

Hasil analisis kemudian diimplementasikan ke dalam dashboard interaktif menggunakan Tableau, seperti yang ditunjukkan pada Gambar 2.20.

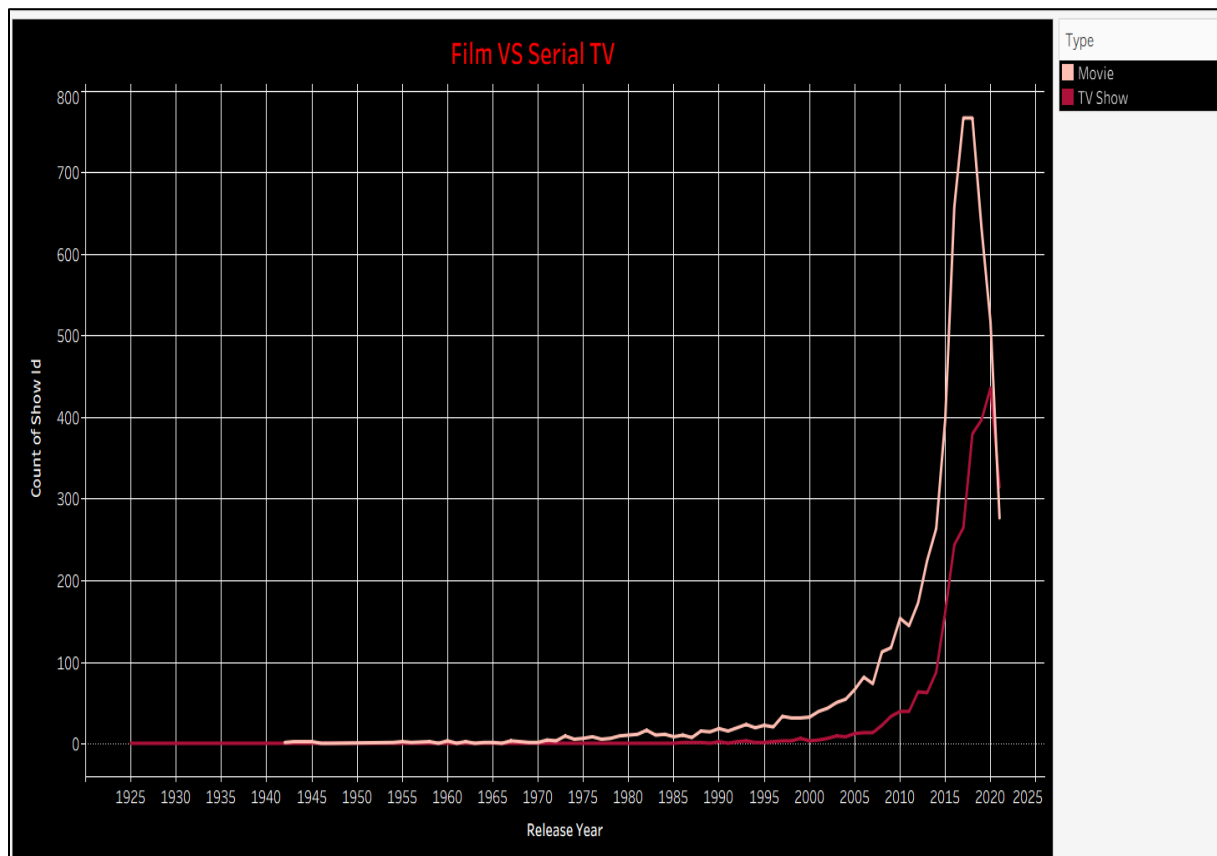
Bab 3

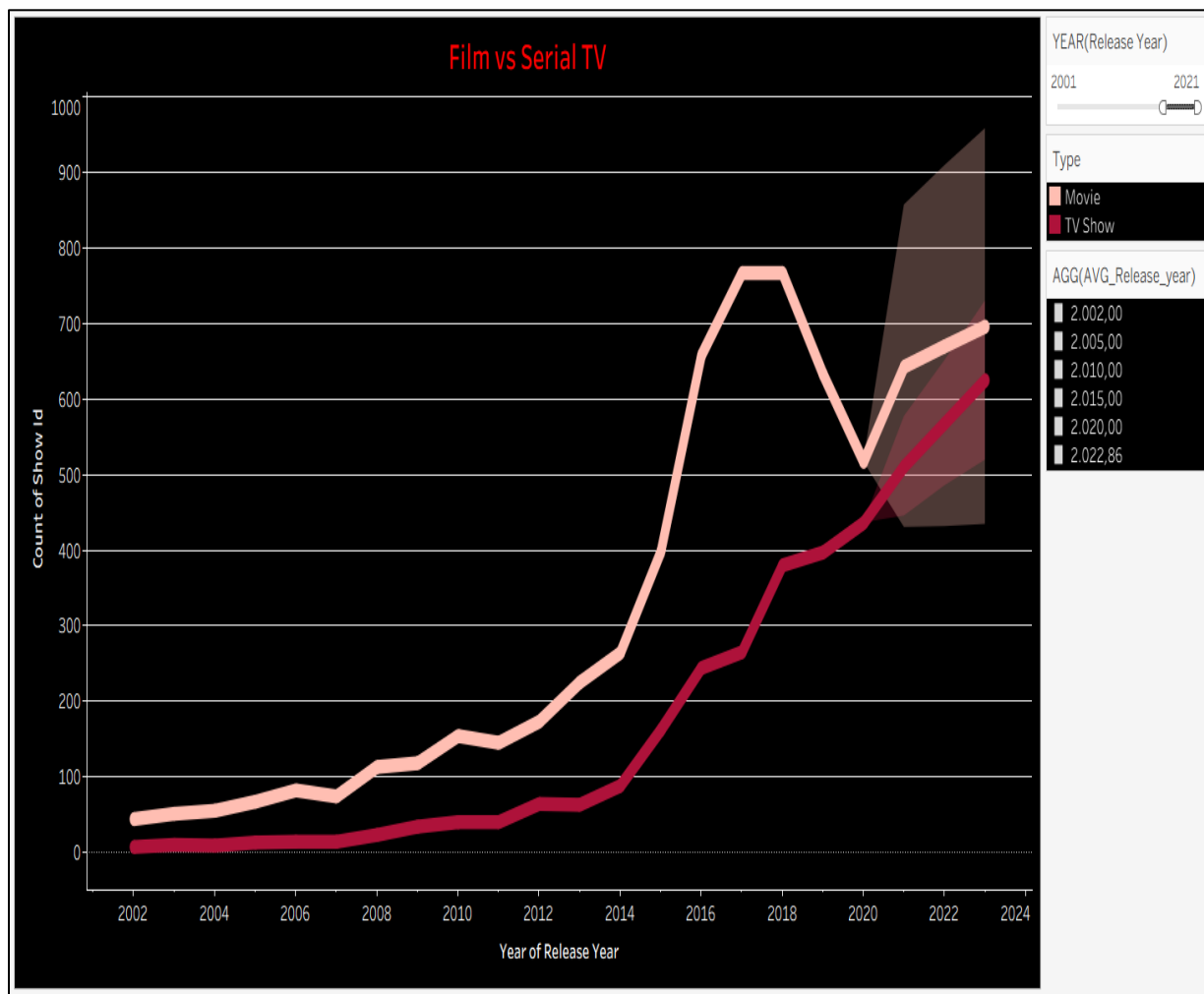
Hasil eksplorasi dan analisis

Bab ini menyajikan hasil eksplorasi dan analisis data dari dataset Netflix Movies and TV Shows yang telah melalui tahap Business Understanding, Data Preparation, Modeling, dan Deployment. Analisis dilakukan secara deskriptif menggunakan Tableau untuk menampilkan berbagai visualisasi interaktif yang menggambarkan pola, tren, dan distribusi konten Netflix berdasarkan jenis tayangan, negara, genre, serta rating penonton.

Hasil dari visualisasi ini digunakan untuk menjawab pertanyaan analitik yang telah dirumuskan pada tahap Business Understanding, serta untuk memperoleh insight yang dapat mendukung pengambilan keputusan bisnis dalam pengelolaan konten di masa mendatang.

3.1 Komposisi Konten: Film vs TV Show

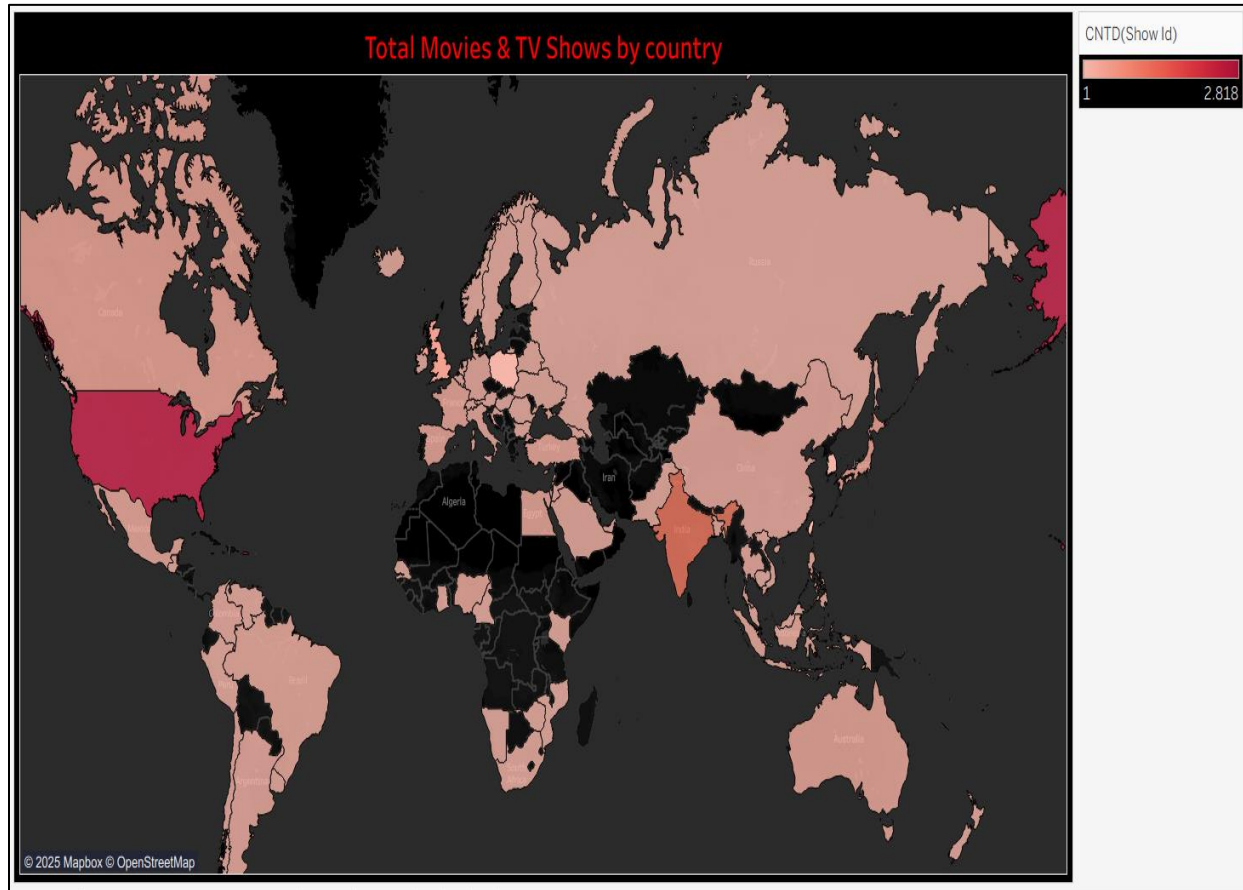




Gambar 3. 1 Perbandingan jumlah Film dan TV Show di Netflix

Gambar 3.1 memperlihatkan perbandingan jumlah konten antara film (Movie) dan acara TV (TV Show) di platform Netflix. Berdasarkan hasil analisis, film mendominasi dengan jumlah 6.131 judul (sekitar 69,62%), sedangkan acara TV berjumlah 2.676 judul (30,38%). Temuan ini menunjukkan bahwa Netflix masih berfokus pada produksi dan distribusi film dibandingkan dengan serial televisi. Untuk visualisasi yang kedua terdapat rata rata tahun release yang dimana rata rata ini berfungsi menunjukkan rata rata tahun berapa yang paling banyak di merilis film maupun Serial TV. Berdasarkan filter tahun yang sudah di setting contohnya seperti yang tertera pada gambar ketika filter ditunjukkan untuk tahun 2001-2021 maka rata rata tahun rilis untuk periode itu adalah 2002, 2005, 2010, 2015, dan 2020.

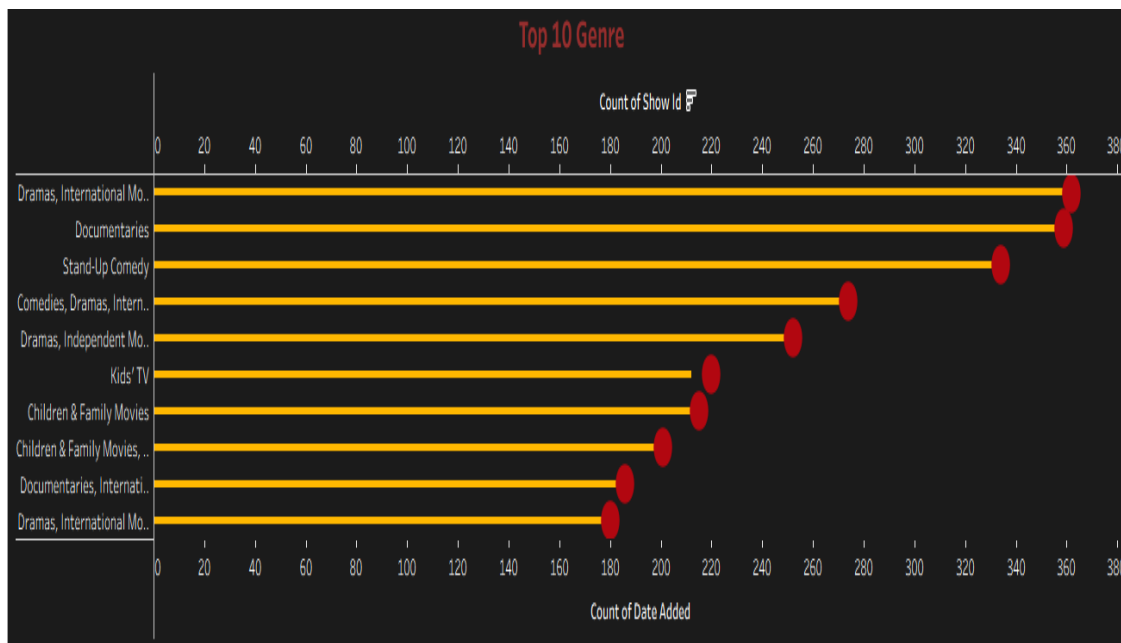
3.2 Distribusi Konten Berdasarkan Negara



Gambar 3. 2 Total Movies & TV Shows berdasarkan Negara

Gambar 3.2 menampilkan distribusi jumlah film dan acara TV berdasarkan negara asal produksinya. Visualisasi menunjukkan bahwa Amerika Serikat merupakan negara dengan jumlah konten terbanyak (sekitar 2.900 judul), diikuti oleh India, Kanada, Jepang, dan Inggris. Dominasi Amerika Serikat menandakan bahwa Netflix masih bergantung kuat pada pasar produksi konten barat, sementara negara-negara Asia mulai menunjukkan kontribusi yang meningkat dalam beberapa tahun terakhir.

3.3 Distribusi Genre Terpopuler

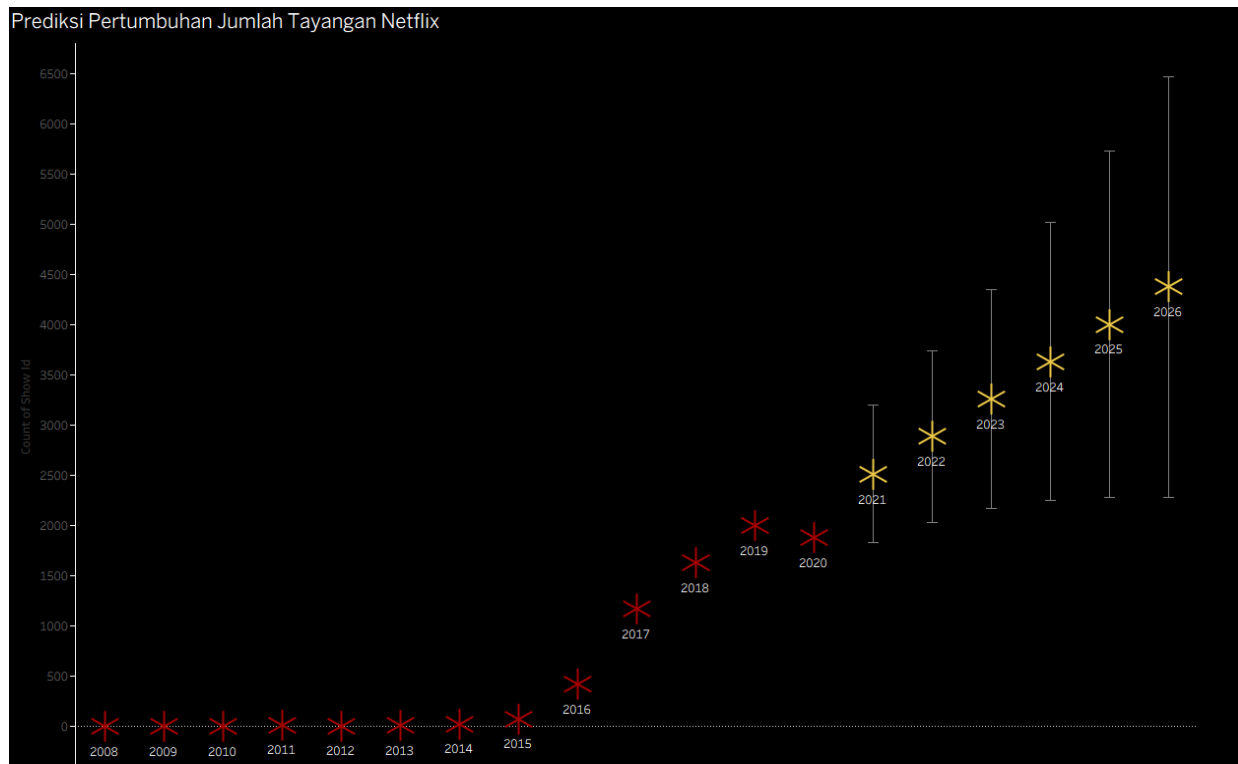


Gambar 3. 3 Top 10 Genre Terpopuler di Netflix

Gambar 3.3 memperlihatkan Top 10 Genre dengan jumlah tayangan terbanyak di Netflix. Berdasarkan hasil visualisasi, genre Drama, International Movies, dan Documentaries menjadi tiga

kategori dengan jumlah penayangan tertinggi. Hal ini menunjukkan bahwa Netflix menempatkan fokus besar pada konten yang bersifat naratif dan mendalam, serta konten dokumenter yang edukatif. Selain itu, titik-titik merah pada grafik menunjukkan aktivitas penambahan tayangan baru genre dengan posisi titik lebih ke ujung batang menandakan peningkatan penambahan konten baru secara konsisten.

3.4 Tren Penambahan Konten per Tahun



Gambar 3. 4 Tren Penambahan Film dan TV Show di Netflix

Gambar 3.4 menampilkan tren pertumbuhan jumlah film dan acara TV dari tahun ke tahun. Terlihat bahwa jumlah konten baru meningkat tajam mulai tahun 2016 hingga mencapai puncaknya pada periode 2017–2019. Setelah itu, laju pertumbuhan cenderung melambat, yang kemungkinan disebabkan oleh perubahan strategi produksi atau dampak pandemi global. Temuan ini menunjukkan bahwa Netflix sempat berfokus pada ekspansi jumlah konten secara besar-besaran, namun kini mulai beralih ke strategi yang lebih menekankan pada kualitas dibanding kuantitas.

3.5 Kesimpulan Hasil Analisis

Berdasarkan hasil analisis data Netflix Film dan TV Show, dapat disimpulkan jika jumlah film dan serial tv selalu meningkat sejak tahun 2008. Pertumbuhan terbesar terjadi pada tahun 2019 dimana terdapat 1999 konten. Hasil prediksi juga menunjukkan jika jumlah judul akan terus bertambah hingga tahun 2025.

Konten Netflix Sebagian besar berasal dari negara Amerika Serikat dan negara lain yaitu India, Inggris dan Kanada. Pada Netflix, konten paling banyak adalah film, meskipun serial tv terus bertambah tiap tahun. Berdasarkan rating TV-MA, mayoritas penonton adalah orang dewasa.

Platform Netflix terus memperluass koleksi dan berbagai tipe konten dari berbagai negara. Hal tersebut menunjukkan jika Netflix fokus ke hal yang membuat penonton tetap tertarik dan juga untuk mempertahankan sebagai salah satu platform streaming terbesar di dunia.

Daftar Pustaka

- [1] A. Ali, A. Raza, M. M. M. Sayed, B. A. Qureshi, and Y. M. Memon, “Data-driven Insights Machine Learning Approaches for Netflix Content Analysis and Visualization,” *J. Eng. Res. Reports*, vol. 27, no. 4, pp. 278–290, 2025, doi: 10.9734/jerr/2025/v27i41471.
- [2] Y. Suhandi, I. Kurniati, and S. Norma, “Penerapan Metode Crisp-DM Dengan Algoritma K-Means Clustering Untuk Segmentasi Mahasiswa Berdasarkan Kualitas Akademik,” *J. Teknol. Inform. dan Komput.*, vol. 6, no. 2, pp. 12–20, 2020, doi: 10.37012/jtik.v6i2.299.
- [3] A. A. A. Fernandes, M. Koehler, N. Konstantinou, P. Pankin, N. W. Paton, and R. Sakellariou, “Data Preparation: A Technological Perspective and Review,” *SN Comput. Sci.*, vol. 4, no. 4, pp. 1–20, 2023, doi: 10.1007/s42979-023-01828-8.