

PROGRAM STUDI SARJANA INFORMATIKA
PROPOSAL PROYEK VISUALISASI DATA



Diajukan Oleh :

Anggota : 1. 11S19003 – Rafelli Simangunsong
2. 11S19007 – Tesalonika Siahaan
3. 11S19025 – Prawita Dwi Friskila
4. 11S19036 – Wybren Manik
5. 11S19046 – Deny Pane

Analisis Visualisasi Netflix Movies and TV Shows (Responses)

FAKULTAS INFORMATIKA DAN TEKNIK ELEKTRO
INSTITUT TEKNOLOGI DEL LAGUBOTI

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Revolusi pada saat ini membawa dampak besar pada industri jasa perfilman. Sebelum adanya revolusi industri tayangan film hanya ditampilkan melalui bioskop, pelanggan harus datang langsung ke bioskop, mengantri dan membeli tiket, lalu menikmati film di ruangan bersama puluhan penonton lain. Bagi beberapa kalangan cara ini dianggap kurang praktis, karena pelanggan yang ingin menonton film harus menyesuaikan waktu dengan jam tayang film yang bersangkutan. Namun saat ini telah hadir layanan Video on Demand (VoD) yakni website penyedia film online secara resmi. Dengan adanya situs VoD penikmat film tidak perlu datang ke bioskop, cukup melakukan aktivasi penonton sudah bisa menikmati film kesukaannya di rumah masing-masing, sesuai dengan jam yang diinginkan, dan tidak terikat pada jam tayang yang ditentukan bioskop.

Salah satu situs web penyedia layanan ini adalah Netflix. Dari situs Kompas.com Netflix merupakan layanan penyedia langganan sewa film, streaming film, program televisi, dan beberapa program yang dibuat oleh Netflix sendiri. Program langganan Netflix memungkinkan pengguna menonton layanan kesukaan dimanapun, kapanpun, dengan media apapun (gadget, smart tv, computer, dan laptop). Netflix seperti toko penyewa DVD yang menawarkan film digital di dunia maya dan resmi. Netflix bersih dari iklan dan tidak ada jadwal penayangan yang ditentukan seperti di TV kabel.

Netflix sudah sangat sering masuk sebagai nominasi dalam ajang penghargaan dan sudah banyak mendapatkan penghargaan. Dikutip dari situs App Annie, Netflix masuk ke dalam peringkat pertama pada Top 10 Apps by All-Time Consumer Spend (2010-2019). Dari tahun ke tahun, pelanggan Netflix di Indonesia selalu menunjukkan peningkatan yang signifikan. Pada 2017, terdapat 95 ribu pelanggan jumlah pelanggan Netflix di Indonesia. Kemudian pada tahun 2018, meningkat hingga 2,5 kali lipat menjadi sebanyak 237,3 ribu. Selanjutnya pada 2019 meningkat menjadi 482 ribu, dan tahun 2020 sudah mencapai 900 ribu lebih pelanggan Netflix di Indonesia. Dilansir dari www.statista.com, dari survei yang telah dilakukan di Amerika Serikat pada Maret 2020 yaitu sebanyak 65% pengguna Netflix berasal dari usia 18-29 tahun. Dengan ini, Netflix lebih populer dan banyak digunakan dikalangan anak muda.

Meski menjadi aplikasi terlaris, Netflix banyak dikeluhkan oleh penggunanya. Beberapa hal yang dikeluhkan adalah setiap kali mengakses layanan netflix, sistem rekomendasi dari si netflix memberikan rekomendasi acara atau film yang terkadang kurang sesuai dengan selera yang paling sering diminati oleh

si orang tersebut. Misalnya untuk anak dari usia 18-29 tahun lebih memilih selera film dengan genre ke percintaan, namun untuk usia diatas 30 tahun lebih banyak memilih genre film komedi. Disini kami akan menampilkan penyebaran *director* Movies and TV Shows dari berbagai negara, perkembangan Movies and TV Shows tiap tahunnya, genre yang ada di data netflix kami ambil yang paling banyak diminati oleh pengguna, kategori *top ratings* Movies and TV Shows pada netflix. Maka dalam proyek ini kelompok kami akan merancang *dashboard* yang di dalamnya terdapat visualisasi data *textual*, *temporal*, dan *spatial* dengan *tableau* sebagai *toolsnya*.

1.2 Tujuan

Tujuan yang diharapkan dalam pegerjaan proyek ini antara lain :

1. Ajukan pertanyaan yang memandu visualisasi data sesuai target konsumen.
2. Lakukan analisis menggunakan kerangka what-why-how untuk menentukan bahasa visualisasi yang akan digunakan.
3. Desain panel yang merupakan visualisasi data teks, waktu dan lokasi.

1.3 Manfaat

Manfaat yang diharapkan dari pengerjaan proyek ini adalah:

1. Bagi mahasiswa, memberikan wawasan serta pengetahuan dalam menerapkan visualisasi data menggunakan tools Tableau. Serta mampu menerapkan aturan-aturan dalam menerapkan visualisasi data.
2. Bagi industri, visualisasi data dapat mendukung pengambilan keputusan dengan adanya analitik untuk memahami potensi pasar, mengukur pendapatan, perbandingan pelanggan dan pesaing.

1.4 Ruang Lingkup

Ruang lingkup pada Netflix Visualization adalah sebagai berikut:

1. Menggunakan *dataset* Netflix Movies and TV Shows.
2. Menggunakan *Tableau Public* sebagai *tool* dalam membangun *dashboard*.

1.5 Pertanyaan Analisis

Tuliskan pertanyaan analisis yang sesuai dengan masalah yang Anda kemukakan pada Latar Belakang berdasarkan kebutuhan audiens.

1. Bagaimana penyebaran *director* Movies and TV Shows dari masing-masing negara?
2. Bagaimana perkembangan Movies and TV Shows pada setiap tahunnya?
3. Apa saja Genre Movies atau TV Shows yang populer pada Netflix?
4. Kategori apa saja yang menduduki 10 *top ratings* Movies and TV Shows pada Netflix?

BAB 2

ISI

Bab isi berisi penjelasan terkait aktivitas pengembangan *dashboard* visualisasi data dalam bentuk alir. Penjelasan meliputi tahap pengumpulan data, analisis data, dan evaluasi visualisasi disertai penjabaran pada setiap tahapannya.

2.1 Analisis

Pada subbab ini berisi penjelasan terkait data dan analisis *What-Why-How Framework* untuk setiap panel visualisasi.

2.1.1 Analisis Data

Pada analisis data dilakukan eksplorasi data dengan tujuan *data understanding* sehingga konteks data, variabel-variabel data dan keterkaitan antar variabel data dapat dipahami dengan baik. *Dataset* yang digunakan yaitu *netflix_titles.csv*. *Dataset* tabular ini berisikan informasi film dan acara tv yang tersedia di Netflix disertai dengan detail seperti informasi pemeran, sutradara, *ratings*, tahun rilis, durasi dan informasi terkait lainnya. *Dataset* terdiri atas 12 atribut/variabel dan 8807 *record*, 8 atribut *nominal*, 3 atribut *ordered ordinal*, dan 1 atribut *ordered quantitative*. Berikut adalah penjelasan dari setiap atribut atau variabel pada *dataset*.

Tabel 1 Informasi Dataset

No.	Nama Atribut	Tipe Atribut	Deskripsi
1	show_id	Nominal	Memberikan informasi tentang id <i>show</i>
2	type	Nominal	Memberikan informasi 2 nilai unik yang berbeda satu adalah Acara TV dan lainnya adalah Film

3	title	Nominal	Memberikan informasi judul Film atau Acara TV
4	director	Nominal	Memberikan informasi sutradara yang mengarahkan Film atau Acara TV
5	cast	Nominal	Memberikan informasi pemeran yang berperan dalam Film atau Acara TV
6	country	Nominal	Memberikan informasi negara rilisnya Film atau Acara TV
7	date_added	Ordered Ordinal	Memberikan informasi tanggal Film atau Acara TV ditambahkan
8	release_year	Ordered Ordinal	Memberikan informasi tahun ketika Film atau Acara TV dirilis
9	rating	Ordered Ordinal	Memberikan informasi Film atau Acara TV yang termasuk dalam kategori mana (misalnya seperti film hanya untuk pelajar, atau orang dewasa, dll)
10	duration	Ordered Quantitative	Memberikan informasi durasi Film atau Acara TV

11	listed_in	Nominal	Memberikan informasi genre Film atau Acara TV
12	description	Nominal	Memberikan informasi deskripsi Film atau Acara TV

2.1.2 Analisis What-Why-How Framework

Pada subbab ini akan dijelaskan analisis mengenai *What-Why-How Framework* untuk setiap panel visualisasi.

2.1.2.1 Analisis What-Why-How Framework Panel Visualisasi 1

Task : Bagaimana penyebaran *director* Movies and TV Shows dari masing-masing negara?

a. What data is shown in the views?

Pertanyaan *what* digunakan untuk menganalisis penyebaran *director* Movies dan TV Shows seperti *attribute type*, *data types*, *dataset availability*, dan *attribute type*. Tujuan analisis yaitu memberikan pemahaman terhadap data dan informasi yang hendak divisualisasikan dalam bentuk *spatial visualization*.

Tabel 2 Analisis What pada Visualisasi 1

No.	<i>Data Abstraction</i>	<i>Type</i>
1	<i>Data Types</i>	<i>Items dan Attribute</i>
2	<i>Dataset Types</i>	<i>Table</i>

3	<i>Dataset Availability</i>	<i>Static (Offline)</i>
4	<i>Attribute Type</i>	<i>Country, Director (Categorical/Nominal)</i> <i>Latitude, Longitude (Generated) (Ordered)</i>

b. Why is the task being performed?

Pertanyaan *why* digunakan untuk menentukan *action* dan *target* dalam abstraksi data.

Tabel 3 Analisis Why pada Visualisasi 1

<i>Action</i>	<i>Target</i>
<ul style="list-style-type: none"> <i>Action → Present → Locate → Identify:</i> <p>Analisis penyebaran <i>director</i> di masing-masing negara diperoleh dengan action untuk menghasilkan visualisasi baru dari <i>data table</i> menjadi <i>data spatial</i>, dengan menggunakan <i>present</i> (menyajikan visualisasi untuk komunikasi dan informasi), <i>locate</i> (mencari tahu di mana objek negara berada), dan <i>identify</i> (menyorot item negara yang ingin diidentifikasi).</p>	<ul style="list-style-type: none"> <i>Target → Spatial Data Shape:</i> <p>Menemukan penyebaran data di seluruh rentang.</p>

c. How is the visualization constructed?

Pertanyaan *how* digunakan untuk menangani kompleksitas data dengan cara berikut:

- 1) *Encode* → *Arrange* → *Separate*: Menyusun data secara spasial.
- 2) *Manipulate* → *Select*: Menyoroti satu atau beberapa item tertentu.

Tahap ini mencakup penentuan idiom visualisasi yang sesuai untuk menjawab pertanyaan berdasarkan analisis *What-Why-How Framework* yang telah dilakukan sebelumnya. Idiom visualisasi dapat menyajikan informasi yang menjawab “Bagaimana penyebaran *director* Movies and TV Shows dari masing-masing negara?” dengan bentuk visualisasi *spatial*. Dengan data *spatial* pembaca dapat melihat penyebaran *director* di masing-masing negara dan menyorot negara bagian yang ingin diidentifikasi.

2.1.2.2 Analisis What-Why-How Framework Panel Visualisasi 2

Task : Bagaimana perkembangan Movies and TV Shows pada setiap tahunnya?

a. What data is shown in the views?

Tabel 4 Analisis What pada Visualisasi 2

No.	<i>Data Abstraction</i>	<i>Type</i>
1	<i>Data Types</i>	<i>Items dan Attribute</i>
2	<i>Dataset Types</i>	<i>Table</i>
3	<i>Dataset Availability</i>	<i>Static (Offline)</i>
4	<i>Attribute Type</i>	Type (<i>Categorical/Nominal</i>)

		Date Added (<i>Ordered</i>)
--	--	-------------------------------

b. Why is the task being performed?

Tabel 5 Analisis Why pada Visualisasi 2

<i>Action</i>	<i>Target</i>
<ul style="list-style-type: none"> • <i>Analyze → Consume → Discover:</i> Analisis dan konsumsi jumlah Movies and TV Shows di setiap tahunnya dan menemukan informasi perkembangan Movies and TV Shows di setiap tahunnya. • <i>Search → Lookup:</i> Pembaca mengetahui informasi apa yang sedang dicari. • <i>Query → Compare:</i> Pembaca melakukan perbandingan dengan mengacu pada beberapa target. 	<ul style="list-style-type: none"> • <i>Target → Trends:</i> Menemukan trends, atau karakteristik tingkat tinggi pada sebuah pola dalam data. • <i>Target → Extremes:</i> Menemukan extremes value, atau distribusi nilai maksimum atau nilai minimum (extremes value) pada data.

c. How is the visualization constructed?

Pertanyaan *how* digunakan untuk menangani kompleksitas data dengan cara berikut:

- 1) *Encode → Arrange → Align*: Menyusun data secara spasial.
- 2) *Manipulate → Select*: Menyoroti satu atau beberapa item tertentu.

Tahap ini mencakup penentuan idiom visualisasi yang sesuai untuk menjawab pertanyaan berdasarkan analisis *What-Why-How Framework* yang telah dilakukan sebelumnya. Idiom visualisasi dapat menyajikan informasi yang menjawab “Bagaimana perkembangan Movies and TV Shows pada setiap tahunnya?” dengan bentuk visualisasi *multiple line chart* dan *heatmap chart*. Dengan demikian pembaca dapat melihat perkembangan Movies and TV Shows pada setiap tahunnya.

2.1.2.3 Analisis What-Why-How Framework Panel Visualisasi 3

Task : Apa saja Genre Movies atau TV Shows yang populer pada Netflix?

a. What data is shown in the views?

Tabel 6 Analisis What pada Visualisasi 3

No.	<i>Data Abstraction</i>	<i>Type</i>
1	<i>Data Types</i>	<i>Items dan Attribute</i>
2	<i>Dataset Types</i>	<i>Table</i>
3	<i>Dataset Availability</i>	<i>Static (Offline)</i>
4	<i>Attribute Type</i>	<i>Genre (Categorical/Nominal)</i>

b. Why is the task being performed?

Tabel 7 Analisis Why pada Visualisasi 3

<i>Action</i>	<i>Target</i>
<ul style="list-style-type: none"> • <i>Action → Present → Locate → Identify:</i> <p>Analisis tren genre yang diperoleh dengan action (menghasilkan visualisasi baru dari data tabel menjadi data <i>textual</i>), <i>present</i> (menyajikan visualisasi untuk komunikasi dan informasi), <i>locate</i>, <i>identify</i> (mengidentifikasi informasi item genre dan rating dengan mengurutkan).</p>	<ul style="list-style-type: none"> • <i>Target → All Data Trends:</i> <p>Menemukan 5 genre movies atau tv shows terbanyak.</p>

c. How is the visualization constructed?

Pertanyaan *how* digunakan untuk menangani kompleksitas data dengan cara berikut:

- 1) *Facet → Filter*: Mengurutkan data.
- 2) *Encode → Arrange → Align*: Menyusun data secara *textual*.

Tahap ini mencakup penentuan idiom visualisasi yang sesuai untuk menjawab pertanyaan berdasarkan analisis *What-Why-How Framework* yang telah dilakukan sebelumnya. Idiom visualisasi dapat menyajikan informasi yang menjawab “Apa saja Genre Movies atau TV Shows yang populer pada Netflix?” dengan bentuk visualisasi *word cloud*. Dengan demikian pembaca dapat melihat apa saja Genre Movies atau TV Shows yang populer pada Netflix dengan melihat warna *word* dan ukuran *word*.

2.1.2.4 Analisis What-Why-How Framework Panel Visualisasi 4

Task : Kategori apa saja yang menduduki 10 *top ratings* Movies dan TV Shows pada Netflix?

a. What data is shown in the views?

Tabel 8 Analisis What pada Visualisasi 4

No.	<i>Data Abstraction</i>	<i>Type</i>
1	<i>Data Types</i>	<i>Items dan Attribute</i>
2	<i>Dataset Types</i>	<i>Table</i>
3	<i>Dataset Availability</i>	<i>Static (Offline)</i>
4	<i>Attribute Type</i>	<i>Title (Categorical/Nominal), Rating (Ordered)</i>

b. Why is the task being performed?

Tabel 9 Analisis Why pada Visualisasi 4

<i>Action</i>	<i>Target</i>
<ul style="list-style-type: none"> <i>Action → Present → Locate → Identify:</i> <p>Analisis tren genre di masing-masing kategori rating diperoleh dengan action (menghasilkan visualisasi baru dari <i>data table</i> menjadi <i>data textual</i>), <i>present</i> (menyajikan visualisasi untuk</p>	<ul style="list-style-type: none"> <i>Target → All Data Trends:</i> <p>Menemukan 10 rating kategori tertinggi.</p>

komunikasi dan informasi), <i>locate</i> , <i>identify</i> (mengidentifikasi informasi item genre dan rating dengan mengurutkan).	
---	--

c. How is the visualization constructed?

Pertanyaan *how* digunakan untuk menangani kompleksitas data dengan cara berikut:

- 1) *Encode* → *Arrange* → *Align*: Menyusun data secara *textual*.
- 2) *Manipulate* → *Select*: Menyoroti satu atau beberapa item tertentu.

Tahap ini mencakup penentuan idiom visualisasi yang sesuai untuk menjawab pertanyaan berdasarkan analisis *What-Why-How Framework* yang telah dilakukan sebelumnya. Idiom visualisasi dapat menyajikan informasi yang menjawab “Kategori apa saja yang menduduki 10 *top ratings* Movies dan TV Shows pada Netflix?” dengan bentuk visualisasi *word bubble*. Dengan demikian pembaca dapat melihat kategori apa saja yang menduduki 10 *top ratings* Movies dan TV Shows pada Netflix dengan melihat *word* dan ukuran *word bubble*.

Set Data

https://drive.google.com/file/d/1HmvyypZvkGSDVBaGNr9FjflQ0Sbt8EaS/view?usp=share_link

Dataset ini mencakup data durasi film Netflix . Pengumpulan data dilakukan dalam kurun 1 tahun terakhir. Kumpulan data ini mencakup durasi film ditonton dan negara mana yang menonton film netflix tersebut.

Referensi

Gunakan Mendeley Desktop dengan *style* IEEE. Sebagai contoh:

- [1] T. Munzner, “Visualization Design and Analysis: Abstractions, Principles, and Methods,” *Department of Computer Science University of British Columbia*, 2012.