

## Analisa Segmentasi Audiens pada Youtube Channel

Chairul Khuldan Hermansyah<sup>1</sup>, Jibril Muhammad Arief<sup>2</sup>, Saimam Abdel<sup>3</sup>, Cahyono Santoso<sup>4</sup>

<sup>1</sup>Sistem Informasi, Universitas Pembangunan Jaya, Tangerang Selatan, Indonesia

<sup>2</sup>Sistem Informasi, Universitas Pembangunan Jaya, Tangerang Selatan, Indonesia

<sup>3</sup>Sistem Informasi, Universitas Pembangunan Jaya, Tangerang Selatan, Indonesia

<sup>4</sup>Sistem Informasi, Universitas Pembangunan Jaya, Tangerang Selatan, Indonesia

*E-mail: cahyono.budy@upj.ac.id*

### Abstrak

Penelitian ini mengkaji segmentasi audiens pada kanal YouTube yang dikelola oleh Neo-Porte, sebuah agensi Virtual YouTuber (VTuber) di Jepang. Tujuannya adalah mengidentifikasi preferensi konten audiens, mengoptimalkan strategi pemasaran, serta memperkuat brand awareness VTuber. Data dikumpulkan dari API YouTube dan scraping manual, lalu dianalisis menggunakan model data warehouse dengan desain tabel fakta dan dimensi untuk mengorganisasi performa video dan interaksi penonton. Implementasi mencakup proses ETL (Extract, Transform, Load) untuk integrasi data, pembuatan skema database relasional, dan dashboard interaktif. Hasil analisis menunjukkan preferensi konten berdasarkan jenis dan waktu publikasi, membantu dalam perencanaan strategi konten yang lebih efektif. Proyek ini mendukung optimalisasi engagement dan loyalitas audiens, peningkatan kolaborasi strategis, serta monetisasi konten yang lebih tinggi. Pendekatan berbasis data ini diharapkan memperkuat posisi Neo-Porte di ekosistem hiburan digital yang kompetitif.

**Kata Kunci:** *Segmentasi Audiens, Virtual YouTuber (VTuber), Data Warehouse*

### Pendahuluan

Pertumbuhan industri hiburan digital telah menciptakan ruang kompetitif bagi konten kreator, termasuk Virtual YouTuber (VTuber). Neo-Porte, sebagai salah satu agensi VTuber terkemuka di Jepang, berfokus pada pengelolaan talenta VTuber dan pembuatan konten digital yang menarik bagi audiens global. Pemahaman yang mendalam tentang segmentasi audiens menjadi penting untuk menyesuaikan strategi konten agar lebih efektif dalam meningkatkan interaksi dan jangkauan pasar (Nugroho, 2024). Rumusan masalah penelitian ini mencakup: Bagaimana preferensi konten audiens dapat diidentifikasi dengan lebih akurat, Bagaimana strategi pemasaran yang lebih efektif dapat dikembangkan berdasarkan data segmentasi audiens, Untuk menjawab pertanyaan ini, penelitian memiliki batasan pada analisis data performa video YouTube VTuber yang diwakili oleh agensi Neo-Porte. Penelitian ini didasarkan pada pengembangan metodologi analitik dengan memanfaatkan data penonton dan performa video yang diambil dari API YouTube dan teknik scraping manual. Studi sebelumnya mengenai analisis data video cenderung terbatas pada evaluasi umum tanpa segmentasi audiens yang mendalam. Kesenjangan ini menunjukkan perlunya pendekatan yang lebih komprehensif untuk memahami pola perilaku dan preferensi penonton (Pratama & Gunawan, 2024).

Penelitian ini bertujuan untuk Mengidentifikasi preferensi konten audiens yang mendukung perancangan konten lebih relevan, Mengoptimalkan strategi pemasaran dan kolaborasi dengan mempertimbangkan kesamaan audiens, Meningkatkan brand awareness VTuber melalui konten yang sesuai dengan karakteristik penonton. Manfaat dari penelitian ini diharapkan mampu memberikan wawasan strategis bagi agensi dalam memanfaatkan data analitik untuk memperkuat posisi di pasar digital. Dari segi kontribusi ilmiah, penelitian ini memberikan pendekatan berbasis data yang dapat menjadi referensi bagi studi lanjutan terkait analisis perilaku penonton dan strategi pemasaran digital di platform video (Role et al., 2012).

Neo-Porte sebagai agensi Virtual YouTuber (VTuber) yang berbasis di Jepang, didirikan untuk membina bakat-bakat VTuber yang menjanjikan dan mengembangkan konten digital yang menarik bagi audiens global. Sebagai salah satu agensi VTuber terkemuka di Jepang, Neo-Porte memiliki misi untuk menciptakan ekosistem hiburan digital yang inovatif dan inklusif, yang dapat dinikmati oleh penggemar di seluruh dunia. Agensi ini mengelola beberapa VTuber populer yang menampilkan berbagai konten, termasuk gameplay, acara langsung (live streaming), interaksi dengan penggemar, serta kolaborasi dengan kreator lainnya (Rakeshkumar Umanath Upadhyay & Prof. Nilesh Choudhary, 2023). Dengan memanfaatkan platform seperti YouTube, Neo-Porte bertujuan untuk mengoptimalkan strategi konten dan pemasaran yang berdasarkan pada data audiens yang tersegmentasi dengan baik. Analisa segmentasi audiens menjadi

## Analisa Segmentasi Audiens pada Youtube Channel

Prosiding Seminar Nasional Vokasi Penerbangan (SNVP) Vol. 3, No. 1, Desember, 2024

penting bagi Neo-Porte untuk memahami preferensi dan perilaku penonton, sehingga dapat membantu dalam merancang konten yang lebih relevan, meningkatkan engagement, serta memperluas jangkauan pasar (Mahendhiran et al., 2024).

### Metode Penelitian

Variabel kunci dalam penelitian ini meliputi total view, rata-rata durasi tonton, total engagement (like dan komentar), dan kategori konten. Penelitian ini bertujuan untuk menganalisis segmentasi audiens pada saluran YouTube yang dikelola oleh agensi Neo-Porte. Materi utama yang digunakan dalam penelitian ini adalah data terkait audiens dan performa video dari saluran YouTube Virtual YouTuber (VTuber). Data ini diperoleh dari dua sumber utama, yaitu API YouTube yang memberikan metrik seperti jumlah tayangan, durasi tonton, jumlah like, dan komentar, serta scraping manual untuk melengkapi data yang tidak tersedia melalui API. Subjek penelitian ini mencakup semua video yang diunggah oleh saluran VTuber Neo-Porte, di mana setiap video dijadikan unit analisis untuk memahami pola keterlibatan audiens. Alat yang digunakan dalam proses penelitian ini adalah skrip Python untuk scraping manual dan sistem manajemen basis data MySQL, yang memungkinkan penyimpanan serta pengelolaan data dalam skema database relasional. Data yang diperoleh melalui API dan scraping disusun dalam database dengan struktur tabel fakta dan tabel dimensi untuk memungkinkan analisis mendalam.

Penelitian ini menggunakan data performa video dalam rentang waktu tiga tahun terakhir, yaitu dari tahun 2022 hingga 2024, dengan pendekatan data warehouse yang memungkinkan analisis data granular pada tingkat video individu. Teknik pengambilan sampel mencakup semua video yang diterbitkan dalam rentang waktu tersebut, sehingga analisis dapat memberikan hasil yang komprehensif mengenai performa saluran. Beberapa variabel kunci yang diukur dalam penelitian ini adalah total view, rata-rata durasi tonton, total engagement (seperti jumlah like dan komentar), serta kategori konten dari setiap video.

Pengumpulan data dilakukan melalui proses ETL (Extract, Transform, Load). Data diekstraksi dari API YouTube dan hasil scraping, kemudian ditransformasikan agar sesuai dengan skema database yang digunakan, dan diunggah ke dalam tabel-tabel yang telah dirancang untuk menyimpan fakta dan dimensi yang diperlukan. Data yang terkumpul kemudian dianalisis dengan metode statistik deskriptif. Visualisasi hasil analisis dilakukan melalui dashboard interaktif yang memberikan tampilan real-time terkait performa video, sehingga analisis dapat mendukung pengambilan keputusan berdasarkan data yang akurat dan terstruktur.

### Pembahasan

Penelitian ini menganalisis segmentasi audiens pada kanal YouTube VTuber yang dikelola oleh agensi Neo-Porte, menggunakan data yang diperoleh dari API YouTube dan teknik scraping manual. Analisis ini memungkinkan pengembangan strategi konten yang lebih efektif, serta identifikasi tren preferensi penonton yang beragam (Babbar et al., 2024). Desain konseptual untuk analisis segmentasi audiens YouTube channel ini berfokus pada bagaimana agensi Neo-Porte dapat menggunakan data analitik untuk mengoptimalkan strategi konten dan pemasaran bagi Virtual YouTuber (VTuber) yang mereka kelola. Dengan menggabungkan pendekatan berbasis data dan kreativitas konten, desain ini bertujuan untuk meningkatkan engagement penonton, memperluas basis audiens, dan memaksimalkan peluang monetisasi. Analisis segmentasi ini akan memberikan insight penting terkait preferensi konten dan pola perilaku audiens, yang semuanya dapat dimanfaatkan untuk pengambilan keputusan yang lebih efektif dalam pengembangan dan promosi konten VTuber (Kwon et al., 2023).

- a. Analisis Preferensi Konten: Menggunakan data penayangan untuk mengetahui jenis konten yang paling diminati (misalnya, gameplay, talkshow, atau live streaming) dan menyesuaikan strategi konten yang lebih relevan.
- b. Strategi Kolaborasi dan Pemasaran: Menggunakan hasil analisis untuk mengidentifikasi peluang kolaborasi dengan VTuber lain atau merek yang memiliki audiens serupa, serta mengembangkan kampanye pemasaran yang lebih efektif.
- c. Pengembangan Brand Awareness VTuber: Membangun dan memperkuat citra VTuber melalui konten yang disesuaikan dengan minat dan karakteristik penonton yang telah tersegmentasi dengan baik.

Desain logis untuk analisa segmentasi audiens YouTube channel ini bertujuan untuk mengorganisasi data secara struktural sehingga memungkinkan analisis yang mendalam dan terperinci mengenai performa kanal VTuber. Dengan menggunakan tabel fakta dan dimensi, desain ini memungkinkan integrasi data video, dan waktu untuk mengidentifikasi pola dan tren yang relevan. Tabel fakta mencatat informasi performa video dan interaksi audiens, sementara tabel dimensi menyediakan konteks tambahan seperti detail video, dan waktu. Struktur ini memfasilitasi pembuatan laporan dan dashboard analitik yang memungkinkan evaluasi kinerja konten dan pengambilan keputusan yang lebih efektif (Akande et al., 2024).

- a. Tabel Fakta

## Analisa Segmentasi Audiens pada Youtube Channel

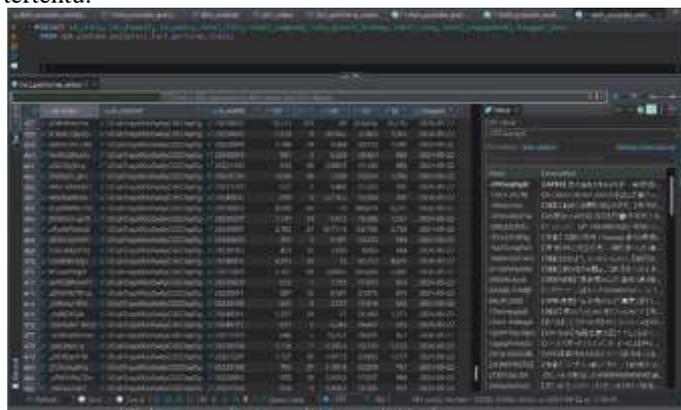
Prosiding Seminar Nasional Vokasi Penerbangan (SNVP) Vol. 3, No. 1, Desember, 2024

1. Fakta Performa Video: Tabel ini menyimpan data performa video pada channel YouTube setiap harinya. Tabel ini mencakup `id_video` (Foreign Key ke dimensi `dim_video`), `id_channel` (Foreign Key ke dimensi `dim_channel`), `id_waktu` (Foreign Key ke dimensi `dim_waktu`), `total_like`, `total_comment`, `rata_durasi_tonton`, `total_view`, `total_engagement`, `tanggal_data`, `id_video`, `id_channel`, `id_waktu`.
- b. Tabel Dimensi
  1. Dimensi Channel (`dim_channel`): Tabel ini menyimpan data mengenai channel YouTube. Atributnya mencakup `id_channel`: ID unik untuk setiap channel, `nama_channel`: Nama dari channel, `total_view`: Jumlah total penayangan seluruh video di channel tersebut, `total_video`: Jumlah total video yang diunggah oleh channel, `total_subscriber`: Jumlah total subscriber dari channel tersebut.
  2. Dimensi Video (`dim_video`): Tabel ini menyimpan data mengenai setiap video yang diunggah. Atributnya mencakup `id_video`: ID unik untuk setiap video, `judul_video`: Judul dari video, `kategori`: Kategori video (seperti Entertainment, Education, dsb.), `durasi`: Durasi video dalam detik, `tanggal_upload`: Tanggal saat video diunggah.
  3. Dimensi Waktu (`dim_waktu`): Tabel ini menyimpan data mengenai waktu terkait performa video. Atributnya mencakup `id_waktu`: ID unik yang merepresentasikan tanggal, `tanggal`: Tanggal data performa dicatat, `bulan`: Bulan ketika data dicatat (1-12), `tahun`: Tahun ketika data dicatat, `hari_dalam_minggu`: Nama hari dalam seminggu.

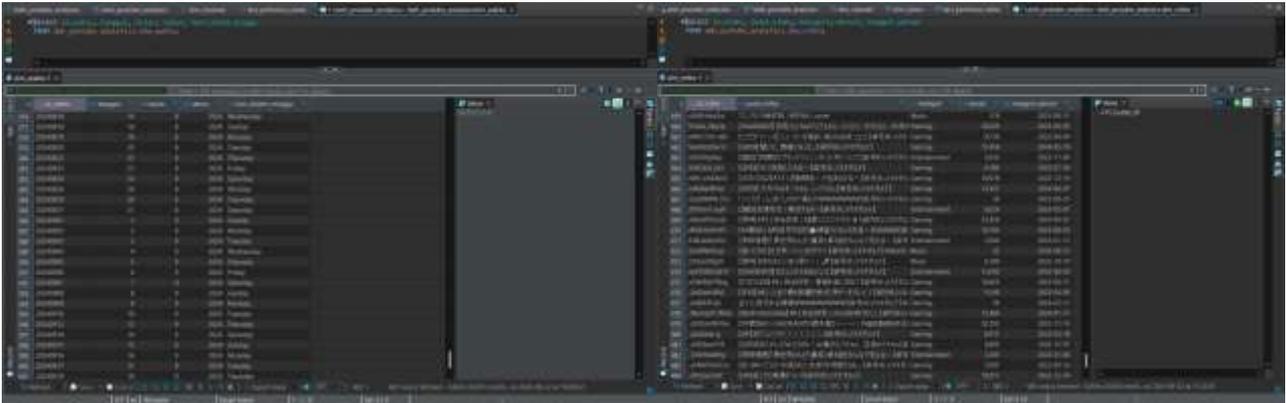
Desain fisik untuk analisa segmentasi audiens YouTube channel mencakup implementasi teknis dari penyimpanan dan pengelolaan data yang diambil melalui API YouTube dan juga scrapping langsung pada kanal YouTube VTuber terkait (Putri et al., 2024). Desain ini memastikan bahwa data yang terkumpul dapat disimpan secara efisien, diakses dengan cepat, dan dianalisis secara efektif untuk mendukung pengambilan keputusan berbasis data (Rakeshkumar Umanath Upadhyay & Prof. Nilesh Choudhary, 2023). Proses ini melibatkan pembangunan struktur database yang efisien, penentuan indeks untuk mempercepat query, proses ETL (Extract, Transform, Load) yang mengintegrasikan data dari berbagai sumber, dan penyimpanan data yang optimal. Tahap ini mencakup:

- a. Membangun Desain Database: Membuat skema database relasional yang mencakup tabel fakta (`fact_performa_video`) dan tabel dimensi (`dim_channel`, `dim_video`, `dim_waktu`). Setiap tabel dirancang dengan kunci primer dan relasi yang jelas untuk memastikan integritas referensial dan konsistensi data. Desain ini akan menggunakan DBMS MySQL.
- b. Menentukan indeks: Menentukan indeks pada kolom-kolom yang sering digunakan dalam kondisi WHERE atau JOIN untuk mempercepat query, seperti `id_video`, `id_channel`, `id_waktu` di tabel fakta, dan kunci utama di setiap tabel dimensi. Indeks ini mengurangi waktu respon query pada saat analisis data, terutama untuk dashboard.
- c. Proses ETL: Mengambil data penonton, performa video, dan metadata dari YouTube menggunakan API YouTube dan juga scrapping langsung pada kanal YouTube VTuber terkait. Proses Extract dilakukan dengan skrip Python atau software pendukung. Data yang diekstrak kemudian di-Transform untuk menyesuaikan dengan skema database, termasuk pembersihan data, normalisasi, dan konversi tipe data. Data yang sudah bersih dan terstruktur kemudian di-Load ke dalam tabel fakta dan dimensi di database.
- d. Penyimpanan data: Data yang telah diolah disimpan di database relasional yang telah dibangun dengan optimalisasi penyimpanan untuk ukuran data yang besar. Penyimpanan menggunakan teknik partisi atau sharding untuk mengelola volume data yang besar secara efisien, serta pengaturan backup dan restore rutin untuk menjaga ketersediaan dan keamanan data.

Capture Screen SQL biasanya merujuk pada tindakan mengambil tangkapan layar (screenshot) dari tampilan hasil query SQL yang dieksekusi di antarmuka perangkat lunak seperti SQL Server Management Studio (SSMS), MySQL Workbench, atau alat pengelola basis data lainnya. Proses ini dilakukan untuk mendokumentasikan hasil query, meninjau data, atau melaporkan hasil tertentu.



**Gambar 1. Table Fact**

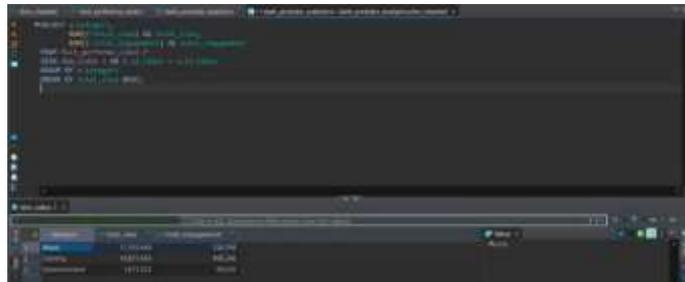


video_id	channel_id	time
1	1	2023-01-01 10:00:00
2	2	2023-01-02 11:00:00
3	3	2023-01-03 12:00:00
4	4	2023-01-04 13:00:00
5	5	2023-01-05 14:00:00
6	6	2023-01-06 15:00:00
7	7	2023-01-07 16:00:00
8	8	2023-01-08 17:00:00
9	9	2023-01-09 18:00:00
10	10	2023-01-10 19:00:00



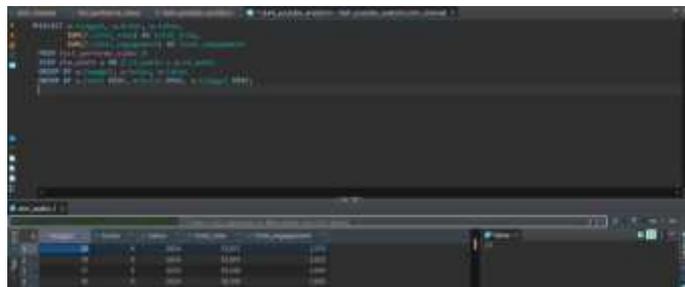
video_id	channel_id	time
1	1	2023-01-01 10:00:00
2	2	2023-01-02 11:00:00
3	3	2023-01-03 12:00:00
4	4	2023-01-04 13:00:00
5	5	2023-01-05 14:00:00
6	6	2023-01-06 15:00:00
7	7	2023-01-07 16:00:00
8	8	2023-01-08 17:00:00
9	9	2023-01-09 18:00:00
10	10	2023-01-10 19:00:00

**Gambar 2. Tabel dim\_waktu, dim\_channel, dim\_video**



video_id	channel_id	time
1	1	2023-01-01 10:00:00
2	2	2023-01-02 11:00:00
3	3	2023-01-03 12:00:00
4	4	2023-01-04 13:00:00
5	5	2023-01-05 14:00:00
6	6	2023-01-06 15:00:00
7	7	2023-01-07 16:00:00
8	8	2023-01-08 17:00:00
9	9	2023-01-09 18:00:00
10	10	2023-01-10 19:00:00

**Gambar 3. Menganalisa Performa Video Berdasarkan Kategori**

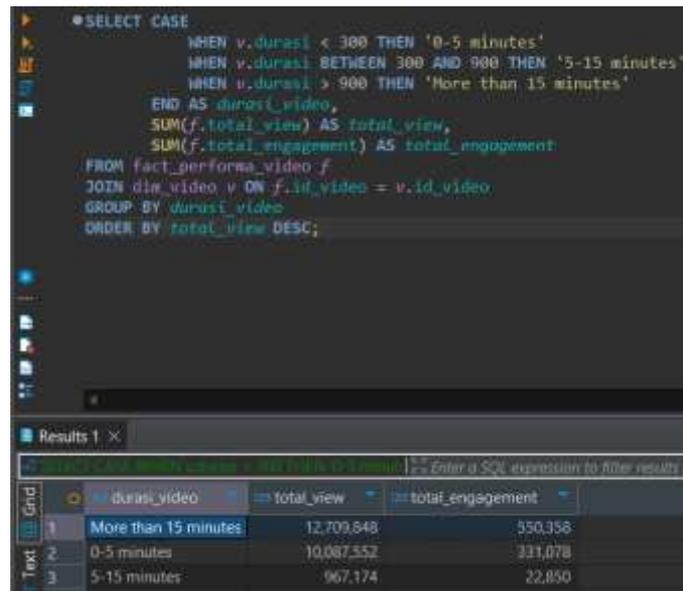


video_id	channel_id	time
1	1	2023-01-01 10:00:00
2	2	2023-01-02 11:00:00
3	3	2023-01-03 12:00:00
4	4	2023-01-04 13:00:00
5	5	2023-01-05 14:00:00
6	6	2023-01-06 15:00:00
7	7	2023-01-07 16:00:00
8	8	2023-01-08 17:00:00
9	9	2023-01-09 18:00:00
10	10	2023-01-10 19:00:00

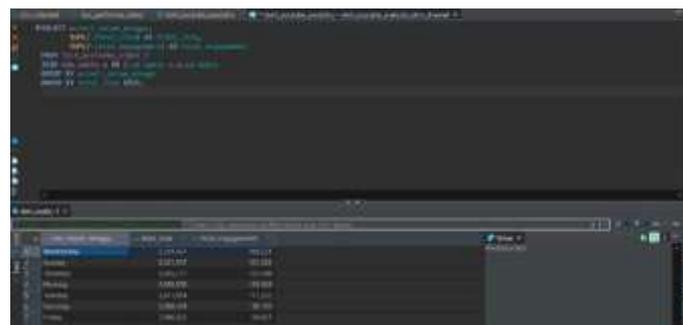
**Gambar 4. Melihat Performa Video Berdasarkan Waktu Upload**

## Analisa Segmentasi Audiens pada Youtube Channel

Prosiding Seminar Nasional Vokasi Penerbangan (SNVP) Vol. 3, No. 1, Desember, 2024



Gambar 5. Melihat Total View dan Total Engagement Berdasarkan Kategori Durasi Video



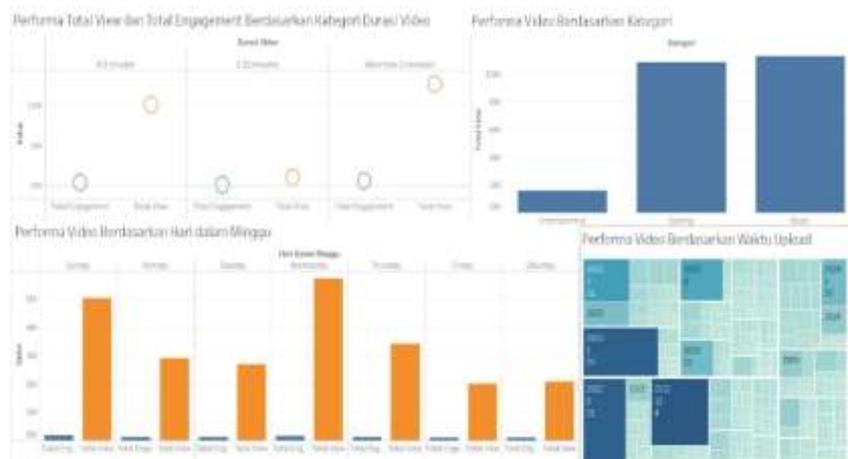
Gambar 6. Melihat Performa Video Berdasarkan Hari dalam Minggu

Dashboard dan reporting untuk analisa segmentasi audiens YouTube channel ini dirancang untuk memberikan tampilan visual yang informatif dan interaktif mengenai performa kanal VTuber. Dengan menggunakan data yang dikumpulkan melalui API YouTube dan scrapping langsung pada kanal YouTube VTuber terkait, dashboard ini memungkinkan pengguna untuk memantau metrik penting seperti total penayangan, rata-rata durasi tonton, dan engagement, serta mengidentifikasi pola dan tren dalam perilaku penonton. Reporting dinamis yang dihasilkan memungkinkan manajer konten dan pemasaran untuk melakukan analisis mendalam dan mengambil keputusan yang tepat guna meningkatkan strategi konten dan pengembangan brand VTuber.

Dashboard Virtual: Dashboard ini dirancang untuk menyajikan visualisasi data yang real-time dan interaktif, seperti grafik garis, pie chart, yang menampilkan matrik utama seperti total view, rata-rata durasi tonton, dan engagement rate. Ini membantu tim manajemen VTuber untuk dengan cepat memahami performa video, melihat perbandingan harian bahkan sampai tahunan, dan mengidentifikasi konten yang paling sukses. Laporan dinamis: memungkinkan pengguna untuk menghasilkan laporan yang dapat disesuaikan berdasarkan kebutuhan analisis. Pengguna dapat memilih berbagai parameter seperti periode waktu. Laporan dinamis ini mendukung strategi berbasis data dengan menyediakan informasi yang dapat diakses dan dianalisis kapan saja. Dashboard dan reporting ini memudahkan agensi dalam memahami audiens dan memaksimalkan strategi konten serta pemasaran, memberikan keunggulan kompetitif dalam dunia VTuber yang semakin kompetitif.

## Analisa Segmentasi Audiens pada Youtube Channel

Prosiding Seminar Nasional Vokasi Penerbangan (SNVP) Vol. 3, No. 1, Desember, 2024



Gambar 6. Capture Screen Dashboard

Performa Total View dan Total Engagement Berdasarkan Kategori Durasi Video pada bagian ini menunjukkan bagaimana perilaku penonton akan membuat perbedaan jumlah total view dan total engagement berdasarkan durasi video, dan grafik ini menampilkan perbedaan antara total view dan engagement berdasarkan durasi dari setiap videonya. Lingkaran biru akan menampilkan total engagement di tiap durasinya dan juga akan menampilkan informasi lebih rinci jumlah total engagement yang telah didapatkan, dan lingkaran oranye juga berfungsi untuk menampilkan total view engagement di tiap durasinya dan juga akan menampilkan informasi rinci dari jumlah total view yang telah didapatkan video tersebut. Diatasnya adalah durasi video yang dipecah menjadi tiga yaitu 0-5 menit, 5-15 menit, dan yang lebih dari 15 menit, dan yang terakhir adalah value merupakan kasaran angka yang telah didapatkan dari kedua lingkaran grafik biru dan oranye. Hal ini akan membantu pengambilan keputusan berdasarkan hasil dari total view dan total engagementnya untuk menentukan durasi video durasi video yang optimal.

Performa Video Berdasarkan Kategori pada grafik ini menunjukkan seberapa baik performa video di berbagai macam kategori yang berbeda dan menunjukkan perbedaan performa antara kategori tersebut. Grafik biru tersebut menunjukkan berbagai macam kategori yang semakin naik berarti total view semakin tinggi dan juga akan menampilkan informasi rinci terkait jumlah total view yang telah didapatkan dari kategori tersebut, selain itu di sebelah kiri juga terdapat total view yang menunjukkan data singkat mengenai jumlah view yang didapatkan per-kategori. Hal ini dapat menjelaskan kategori mana yang lebih disukai oleh penonton dan akan membantu pengambilan keputusan untuk membuat video dengan kategori yang sesuai dengan minat penonton.

Performa Video Berdasarkan Hari dalam Minggu pada grafik ini menunjukkan bagaimana performa video di setiap harinya dan dapat melihat perbedaan antara total engagement yang diwakili dengan grafik batang berwarna biru dan total view yang diwakili dengan grafik batang berwarna oranye. Diatas grafik batang terdapat Hari dalam minggu yang memisahkan performa video di setiap harinya. Grafik ini dapat membantu untuk memperkirakan bagaimana hasil dari performa video tersebut dengan mengacu kepada data mingguan yang telah ditampilkan di grafik. Hal ini akan membantu pengambilan keputusan untuk menyusun jadwal mingguan.

Performa Video Berdasarkan Waktu Upload pada bagian ini merupakan data besar yang berisi performa video dalam 3 tahun kebelakang. Grafik ini berisi performa video yang menunjukkan tahun bulan dan tanggal dibuatnya video tersebut dan tiap persegi juga akan menampilkan informasi lebih lanjut seperti bulan, tahun, tanggal, jumlah total engagement, dan jumlah dari total view yang telah didapatkan dari masing masing grafik yang tidak ter-highlight dikarenakan performanya tidak terlalu tinggi. Grafik ini akan menampilkan konten atau performa video yang dari yang paling baik hingga yang kurang dengan jarak rentan waktu 3 tahun kebelakang. Hal ini dapat membantu dalam melihat perubahan preferensi penonton dan mengevaluasi keberhasilan konten sebelumnya untuk perencanaan konten yang lebih baik.

Nine Steps Methodology adalah kerangka kerja sistematis yang digunakan dalam berbagai bidang, seperti manajemen proyek, implementasi teknologi, perencanaan strategis, atau pemecahan masalah, untuk memastikan semua aspek dari proses dijalankan dengan benar dan efisien. Meskipun detailnya bisa bervariasi tergantung pada konteks, berikut ini adalah gambaran umum dari sembilan langkah yang sering digunakan dalam metodologi ini:

- Choosing the Process yang dipilih dalam analisa ini adalah analisis segmentasi audiens pada YouTube channel VTuber yang dikelola oleh Neo-Porte. Proses ini melibatkan pengumpulan data dari API YouTube serta scrapping untuk memetakan perilaku audiens, preferensi konten, dan performa video yang diunggah. Hasil analisis digunakan untuk meningkatkan engagement dan strategi pemasaran.

## Analisa Segmentasi Audiens pada Youtube Channel

Prosiding Seminar Nasional Vokasi Penerbangan (SNVP) Vol. 3, No. 1, Desember, 2024

- b. Choosing the Grain yang akan digunakan mencakup tingkat video individu. Data ini meliputi performa harian dari setiap video VTuber, seperti jumlah total penayangan, total interaksi (like, comment), dan rata-rata durasi tontonan. Selain itu, granularitas data juga mencakup dimensi waktu (hari, bulan, tahun) serta kategori konten (genre, durasi video).
- c. Identifying and Confirming the Dimensions, dimensi yang digunakan dalam analisis ini meliputi Dimensi Video: Informasi tentang video (judul, kategori, durasi, tanggal upload), Dimensi Channel: Informasi tentang channel (nama channel, jumlah subscriber, total view, total video), Dimensi Waktu: Data terkait waktu (tanggal, bulan, tahun, hari dalam minggu). Konfirmasi dilakukan dengan memastikan setiap dimensi terhubung dengan tabel fakta melalui id foreign key untuk memudahkan analisis performa dan perilaku audiens.
- d. Choosing the Facts yang akan dianalisis dalam data warehouse untuk Analisa segmentasi audiens pada Youtube Channel Neo-Porte mencakup performa video, yang terdiri dari
  1. Fakta Performa Video: Tabel ini menyimpan data performa video pada channel YouTube setiap harinya. Tabel ini mencakup id\_video (Foreign Key ke dimensi dim\_video): ID unik untuk setiap video, id\_channel (Foreign Key ke dimensi dim\_channel): ID unik untuk setiap channel yang mengupload video, id\_waktu (Foreign Key ke dimensi dim\_waktu): ID yang mewakili waktu (tanggal) data performa dicatat, total\_like: Jumlah total like yang diterima video, total\_comment: Jumlah total komentar pada video, rata\_durasi\_tonton: Rata-rata durasi tontonan video dalam detik (waktu video, total view), total\_view: Jumlah total penayangan video, total\_engagement: Jumlah total interaksi video (like, comment), tanggal\_data: Tanggal saat data diambil.
  2. Relasi antar tabel mencakup id\_video: merujuk ke tabel dimensi dim\_video untuk informasi tentang video, id\_channel: merujuk ke tabel dimensi dim\_channel untuk informasi tentang channel yang mengupload video, id\_waktu: merujuk ke tabel dimensi dim\_waktu untuk informasi waktu video.
- e. Storing PreCalculations in The Facts Table Setelah memilih fakta yang akan disimpan, dilakukan pra-perhitungan untuk meningkatkan efisiensi analisis. Perhitungan ini meliputi jumlah total penayangan dan interaksi dari masing-masing video, serta rata-rata durasi tontonan berdasarkan kategori konten. Misalnya, menghitung total engagement video untuk konten gameplay, talkshow, dan live streaming. Ini memastikan bahwa tabel fakta siap untuk digunakan dalam analisis dan pembuatan laporan.
- f. Rounding out the Dimension Table tahap ini adalah langkah untuk melengkapi tabel dimensi yang diperlukan untuk analisis segmentasi audiens YouTube. Fakta performa video akan dihubungkan ke tabel dimensi untuk mendukung analisis yang intuitif dan mudah dimengerti oleh pengguna. Berikut adalah deskripsi tabel dimensi yang relevan:
  1. Dimensi Video (dim\_video) meliputi Key: id\_video, Deskripsi: Dimensi ini berisi informasi tentang setiap video yang diunggah. Hal ini berguna untuk menganalisis performa video secara spesifik berdasarkan judul, kategori, durasi, dan tanggal unggahan, Field:
  2. Dimensi Channel (dim\_channel) meliputi Key: id\_channel, Deskripsi: Dimensi ini berisi informasi mengenai channel YouTube. Hal ini penting untuk mengidentifikasi channel dengan performa terbaik serta membandingkan performa antar channel dan Field:
  3. Dimensi Waktu (dim\_waktu) meliputi Key: id\_waktu, Deskripsi: Dimensi ini menyimpan data tentang waktu terkait dengan performa video. Analisis waktu sangat penting untuk memahami tren berdasarkan hari, bulan, dan tahun dan Field.Tabel-tabel dimensi ini akan digunakan untuk memperkaya analisis performa video berdasarkan berbagai perspektif, seperti kategori video, waktu unggahan, jumlah engagement, dan data channel. Dengan adanya tabel dimensi, pengguna dapat menelusuri data performa video dengan lebih mudah dan intuitif.
- g. Choosing the Duration of the Dimension pada tahap ini, durasi waktu yang digunakan adalah data performa video selama 3 tahun ke belakang (2022 hingga 2024) untuk memahami tren dan perubahan perilaku audiens dari waktu ke waktu. Data ini diambil dari dimensi waktu (dim\_waktu), yang mencakup data tahunan, bulanan, dan mingguan. Durasi ini penting untuk mengevaluasi perubahan tren konten yang diminati oleh audiens.
- h. Tracking Slowly Changing Dimension diman kami menerapkan SCD tipe 1, yaitu dengan menggantikan data lama secara langsung, karena atribut dalam dimensi kami memiliki kunci unik yang diperoleh langsung dari API YouTube

```
SELECT id_channel, nama_channel, total_view, total_video, total_subscriber
FROM dwh_youtube_analytics.dim_channel;
```

id_channel	nama_channel	total_view	total_video	total_subscriber
UCuk7vapXKckSw6jCGGDspDg	#1月ゆい / Hizuki Yui	48,658,941	948	352,000

Gambar 7. Tabel Dimensi Channel

## Analisa Segmentasi Audiens pada Youtube Channel

Prosiding Seminar Nasional Vokasi Penerbangan (SNVP) Vol. 3, No. 1, Desember, 2024



```
SELECT id_video, judul_video, kategori, durasi, tanggal_upload
FROM dw.youtube_analytics.dim_video;
```

id_video	judul_video	kategori	durasi	tanggal_upload
CPC0eqNyl0	【APEX】バハルカスタムLVモーター/流石ハルさん、カクナナさん【新月】Gaming	Gaming	8,299	2023-04-06
DG3r-mUJik	【A Dance of Fire and Ice】激ムズ音ゲーおもしろい!?【新月】Gaming	Gaming	19,716	2024-04-11
-FBleVct3lo	【雑談】起きた時間が朝なので。【新月】Entertainment	Entertainment	7,502	2022-09-16
-h88m582DsQ	【雑談】おはよう、いってらっしゃい【新月】Entertainment	Entertainment	9,798	2024-05-27
-SiYx33HMYg	【生歌】狂歌の花唄 / Vaundy 歌う気持ちも良すぎるぞ【新月】Music	Music	51	2023-05-09

Gambar 8. Tabel Dimensi Video

- i. Deciding The Physical Design ditentukan berdasarkan kebutuhan analisis, seperti memprioritaskan pengambilan data tentang performa video berdasarkan kategori konten, waktu upload, dan total engagement. Pengguna dapat melakukan query secara real-time melalui dashboard yang disediakan, memungkinkan mereka untuk melihat metrik performa utama dan mengambil keputusan strategis berdasarkan hasil analisis.
  1. Membangun Desain Database mencakup pembuatan skema relasional yang terdiri dari tabel fakta dan tabel dimensi : Tabel Fakta: fact\_performa\_video yang mencakup performa video YouTube seperti jumlah like, komentar, durasi tonton, total view, dan total engagement ; Tabel Dimensi: dim\_channel, dim\_video, dan dim\_waktu yang memuat informasi detail terkait video, channel, dan waktu. Setiap tabel dirancang dengan kunci primer dan relasi yang jelas untuk menjaga integritas referensial dan konsistensi data. Dalam implementasinya, sistem ini menggunakan DBMS MySQL yang dioptimalkan untuk kebutuhan analisis data dalam jumlah besar.
  2. Menentukan Indeks Untuk mempercepat query pada proses analisis data, indeks harus ditetapkan pada kolom-kolom kunci yang sering digunakan dalam operasi WHERE atau JOIN. Kolom yang diindeks meliputi: id\_video (pada fact\_performa\_video dan dim\_video); id\_channel (pada dim\_channel); id\_waktu (pada dim\_waktu). Indeks ini akan mempercepat pengambilan data performa video dan interaksi audiens di berbagai dimensi, terutama untuk kebutuhan dashboard.
  3. Proses ETL (Extract, Transform, Load) Tahap ETL merupakan bagian penting dalam manajemen data. Proses ini melibatkan: Extract: Mengambil data penonton dan performa video dari API YouTube serta scrapping langsung dari kanal VTuber terkait menggunakan skrip Python atau perangkat lunak pendukung; Transform: Data yang diambil melalui API kemudian diubah agar sesuai dengan skema database. Proses ini termasuk pembersihan data, normalisasi, dan konversi tipe data untuk memastikan data konsisten dan siap dianalisis; Load: Data yang sudah bersih dan sesuai dengan skema kemudian dimasukkan ke dalam tabel fakta dan tabel dimensi yang telah dirancang.
  4. Penyimpanan Data Untuk memastikan kinerja optimal dalam menyimpan dan mengakses data yang besar, beberapa teknik penyimpanan diterapkan: Partisi atau Sharding: Teknik ini digunakan untuk memecah data menjadi bagian-bagian yang lebih kecil berdasarkan periode waktu atau dimensi tertentu, sehingga mempermudah pengelolaan data besar dan mempercepat proses query; Backup dan Restore: Pengaturan backup dan restore rutin diperlukan untuk menjaga keamanan data dan memastikan ketersediaan data yang andal untuk kebutuhan analisis jangka panjang.

## Kesimpulan

Proyek analisa segmentasi audiens YouTube channel untuk agensi Virtual YouTuber seperti agensi Neo-Porte memberikan wawasan mendalam mengenai perilaku dan preferensi penonton. Dengan memanfaatkan data yang dikumpulkan melalui API YouTube, serta diterapkan melalui desain logis dan fisik database yang terstruktur, proyek ini memungkinkan pengelolaan data yang efisien dan analisis yang komprehensif. Implementasi dashboard virtual memungkinkan tim manajemen untuk memvisualisasikan performa konten secara real-time dan menghasilkan laporan analitik yang dapat disesuaikan, membantu dalam pengambilan keputusan yang lebih cepat dan tepat. Hasil dari analisis ini akan mendukung agensi dalam merancang strategi konten dan pemasaran yang lebih efektif, meningkatkan engagement dan loyalitas penonton, serta mengoptimalkan peluang monetisasi dan kolaborasi dengan VTuber lain atau merek terkait. Dengan pendekatan berbasis data, agensi dapat lebih memahami segmentasi audiens yang beragam dan memaksimalkan potensi kanal YouTube VTuber yang mereka kelola, sehingga menciptakan pertumbuhan yang berkelanjutan dan memperkuat posisi mereka di industri hiburan digital.

## Analisa Segmentasi Audiens pada Youtube Channel

Prosiding Seminar Nasional Vokasi Penerbangan (SNVP) Vol. 3, No. 1, Desember, 2024

### Ucapan Terima Kasih

Penulis mengucapkan terima kasih kepada Instansi/perusahaan/lembaga yang telah memberi dukungan yang membantu pelaksanaan penelitian dan atau penulisan artikel.

### Daftar Pustaka

- Akande, O., Asani, E. O., & Dautare, B. (2024). Customer Segmentation Through RFM Analysis and K-Means Clustering: Leveraging Data-Driven Insights for Effective Marketing Strategy. *Ceddi Journal of Information System and Technology (JST)*, 3(1), 14–25. <https://doi.org/10.56134/jst.v3i1.81>
- Babbar, G., Anand, A., & Agarwal, M. (2024). Selection of Optimal YouTube View Count Prediction Model Using Data Envelopment Analysis Approach. *Journal of Computational and Cognitive Engineering*, 00(July), 1–13. <https://doi.org/10.47852/bonviewjccce42023120>
- Kwon, N., Stewart, R. E., Wang, X., Marzalik, J. S., Bufka, L. F., Halfond, R. W., & Purtle, J. (2023). Where do psychologists turn to inform clinical decisions? Audience segmentation to guide dissemination strategies. *Implementation Research and Practice*, 4. <https://doi.org/10.1177/26334895231185376>
- Mahendhiran, P. D., Manickam, H., Sadhanantham, B., & Duraisamy, S. (2024). *Improving Customer Engagement via Segmentation Empowered by Machine Learning*. 19, 206–194. <https://doi.org/10.29007/wx1m>
- Nugroho, D. A. (2024). Use Of Digital Content As A Means To Reach Unreachable Market Segments Conventionally. *Eduvest - Journal of Universal Studies*, 4(6), 5242–5248. <https://doi.org/10.59188/eduvest.v4i6.1493>
- Pratama, R. H., & Gunawan, P. H. (2024). YouTube Viewership Increation Analysis and Prediction using Facebook Prophet Model. *Jurnal Media Informatika Budidarma*, 8(1), 383. <https://doi.org/10.30865/mib.v8i1.7240>
- Putri, V. R., Mani, L., Fedora, A., Amanina, S., & Aras, M. (2024). The Influence of Digital Content, Public Figure Self-Branding, Celebrity Worship, and Audience Motivation to Increase Audience Interest in the YouTube Starpro Indonesia Social Media Account. *Jurnal Indonesia Sosial Teknologi*, 5(8), 2908–8915. <https://doi.org/10.59141/jist.v5i8.1322>
- Rakeshkumar Umanath Upadhyay, & Prof. Nilesh Choudhary. (2023). A Review of Customer Segmentation Methods. *International Journal of Advanced Research in Science, Communication and Technology*, 3(2), 684–692. <https://doi.org/10.48175/ijarsct-8904>
- Role, T. H. E., Scenarios, O. F., Global, I. N., Studies, C., For, S., Future, A., & Biodiversity, T. O. (2012). SPECIAL IS S U E Scenarios as a tool for large-scale ecological research : experiences and SCENARIOS. *Ecology*, 1–4.