

*Analisis Pengelompokan Lagu Terpopuler Spotify*



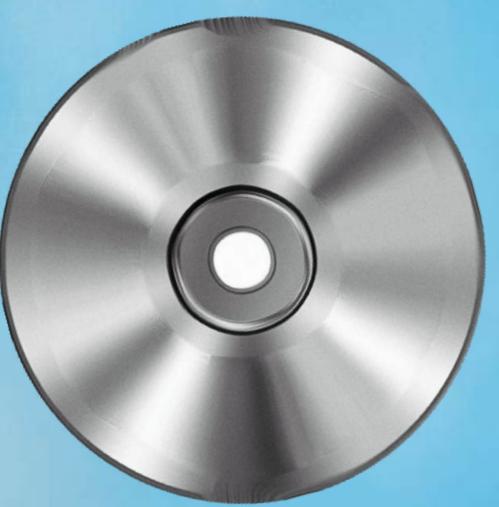
*Menggunakan Algoritma K-Means*

*Berdasarkan Karakteristik Audio dan Popularitas*

oleh :

- Arko Fernanda Wibawa (24782006)
- Juwita Valentiya (24782014)

( Teknologi Rekayasa Internet 3A )



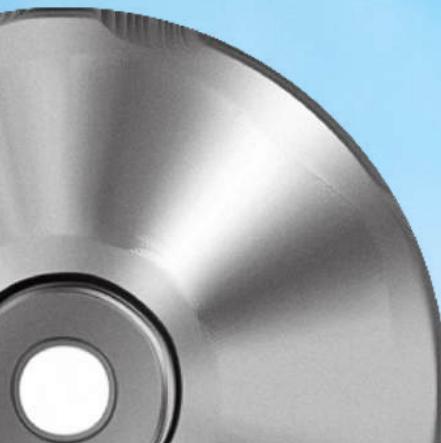
## *Permasalahan Penelitian*

Banyaknya lagu yang tersedia di platform Spotify menyebabkan kesulitan dalam memahami pola popularitas lagu yang sebenarnya, sementara data popularitas seperti total streams dan rata-rata streams harian yang tersedia belum dimanfaatkan secara optimal untuk mengelompokkan lagu terpopuler, sehingga belum diketahui bagaimana pola pengelompokan lagu terpopuler Spotify berdasarkan indikator popularitas streaming menggunakan algoritma K-Means serta jumlah cluster yang paling sesuai dan karakteristik dari setiap kelompok lagu yang terbentuk.



## *Tujuan Penelitian*

Penelitian ini bertujuan untuk menganalisis dan mengelompokkan lagu-lagu terpopuler Spotify berdasarkan indikator popularitas streaming menggunakan algoritma K-Means guna mengetahui pola popularitas lagu, menentukan jumlah cluster yang paling optimal, serta mengidentifikasi karakteristik setiap kelompok lagu yang terbentuk sebagai bahan analisis dan rekomendasi.





# *Analisis Data Eksploratif (Exploratory Data Analysis / EDA)*

Analisis Data Eksploratif (EDA) dilakukan untuk memahami karakteristik awal dataset lagu terpopuler Spotify sebelum dilakukan proses clustering. Pada tahap ini, data dianalisis menggunakan statistik deskriptif dan visualisasi untuk melihat distribusi indikator popularitas, seperti total streams dan rata-rata streams harian, serta hubungan antar variabel yang digunakan. Hasil EDA menunjukkan bahwa distribusi data cenderung tidak merata dan memiliki perbedaan skala yang cukup besar antar lagu, sehingga diperlukan transformasi log dan standardisasi data. Selain itu, analisis korelasi memperlihatkan adanya hubungan antara total streams dan streams harian, yang mengindikasikan bahwa indikator popularitas saling berkaitan dan relevan untuk digunakan sebagai fitur dalam proses pengelompokan lagu terpopuler Spotify.



# Pentingnya Popularitas Lagu



**Menunjukkan minat pendengar**  
Popularitas lagu menunjukkan tingkat ketertarikan dan penerimaan pendengar terhadap sebuah lagu di Spotify.



**Menggambarkan relevansi lagu**  
Total streams dan streams harian membantu mengetahui apakah lagu sedang populer, stabil, atau mulai menurun.



**Dasar pengelompokan lagu**  
Data popularitas digunakan untuk menganalisis pola dan mengelompokkan lagu berdasarkan karakteristik yang serupa.



**MENDUKUNG STRATEGI MUSIK DIGITAL**  
Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua.



# Data Mining & Clustering

- Data Mining
- Data mining merupakan proses pengolahan dan analisis data dalam jumlah besar untuk menemukan pola, informasi, atau pengetahuan yang sebelumnya tidak terlihat.
- Clustering sebagai Unsupervised Learning
- Clustering adalah salah satu teknik data mining yang mengelompokkan data tanpa label, berdasarkan tingkat kemiripan karakteristik yang dimiliki.
- Penerapan pada Lagu Spotify
- Dalam penelitian ini, clustering digunakan untuk mengelompokkan lagu terpopuler Spotify berdasarkan indikator popularitas streaming agar pola popularitas lebih mudah dianalisis.
- Manfaat Clustering
- Hasil clustering membantu memahami segmentasi lagu, sehingga dapat dimanfaatkan untuk analisis tren, rekomendasi, dan strategi musik digital.





# DATASET PENELITIAN



- **Sumber Database**

Dataset yang digunakan berasal dari Spotify Most Streamed Songs yang memuat data lagu-lagu terpopuler di platform Spotify berdasarkan jumlah pemutaran.

- **Jumlah dan Bentuk Data**

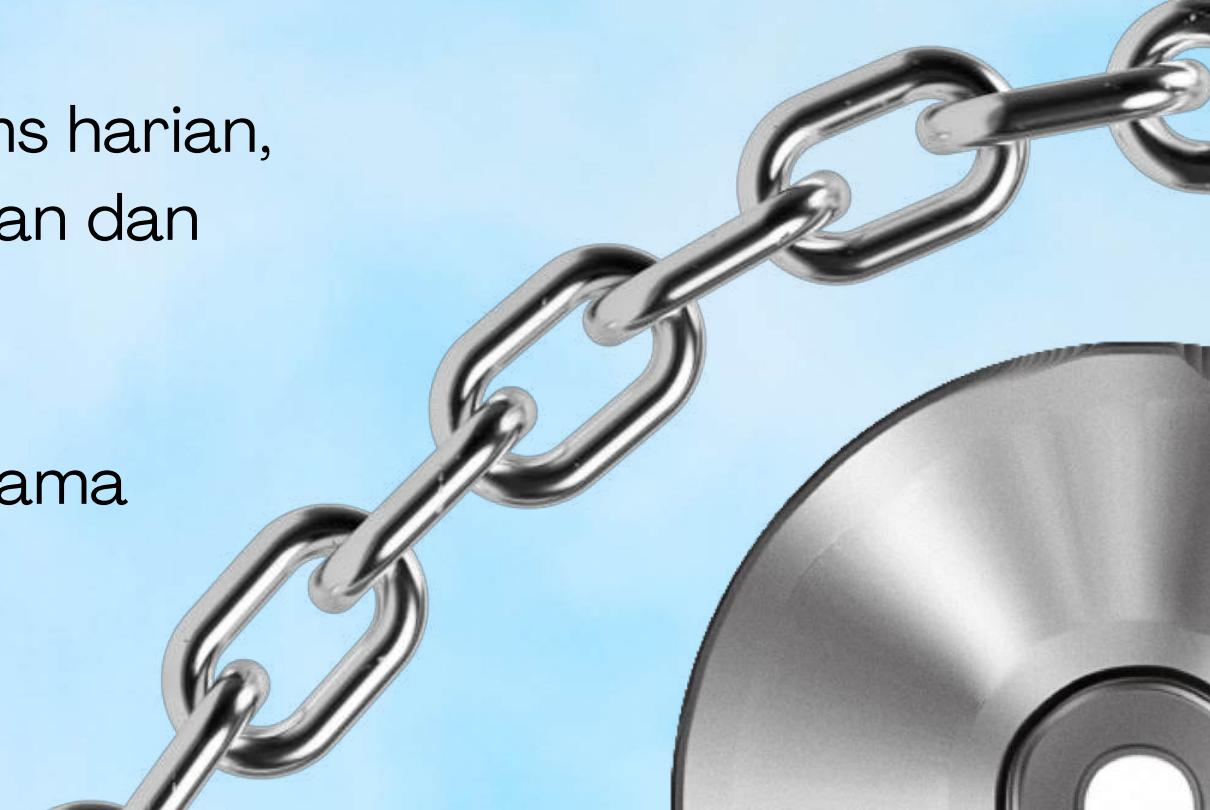
Dataset terdiri dari 2.500 lagu, di mana setiap baris merepresentasikan satu lagu dengan informasi popularitas streaming.

- **Informasi yang Dikandung**

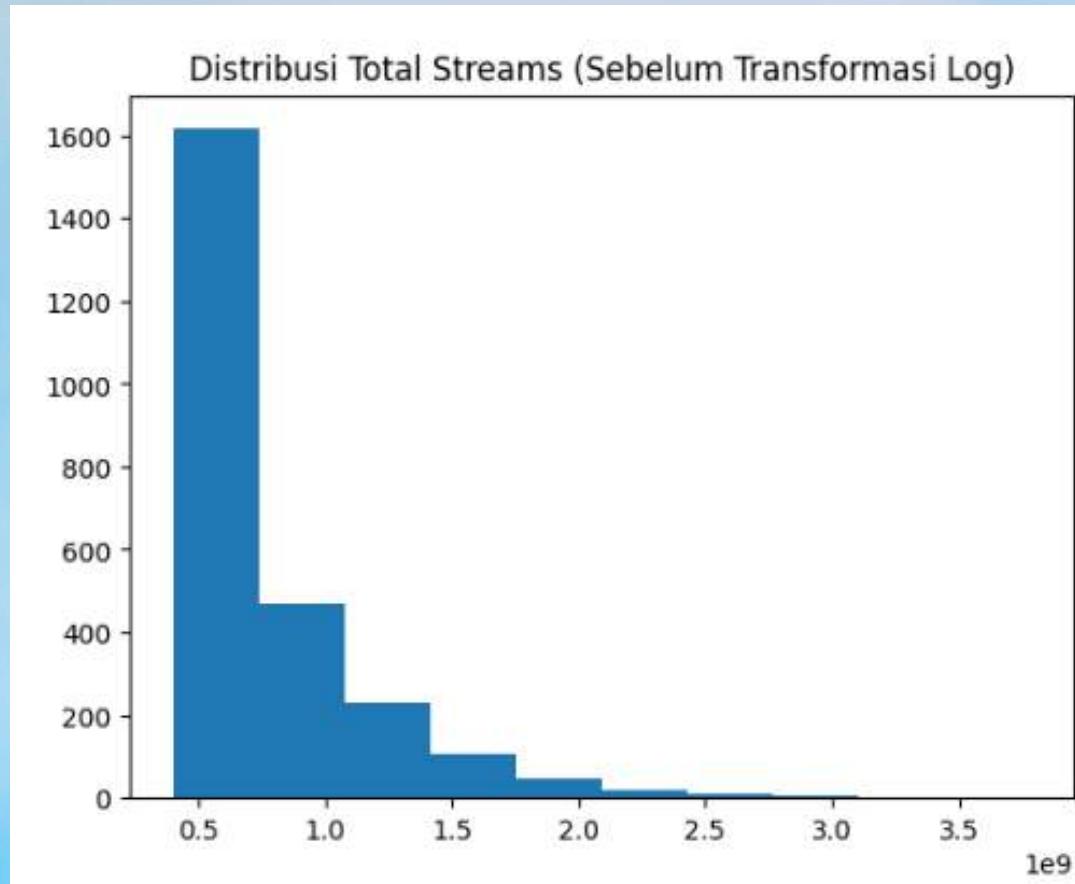
Data mencakup indikator utama popularitas lagu, yaitu total streams dan streams harian, yang digunakan untuk menggambarkan tingkat popularitas lagu secara keseluruhan dan dari waktu ke waktu.

- **Peran Dataset dalam Penelitian**

Dataset ini digunakan pada tahap EDA, preprocessing data, dan sebagai input utama dalam proses clustering menggunakan algoritma K-Means.

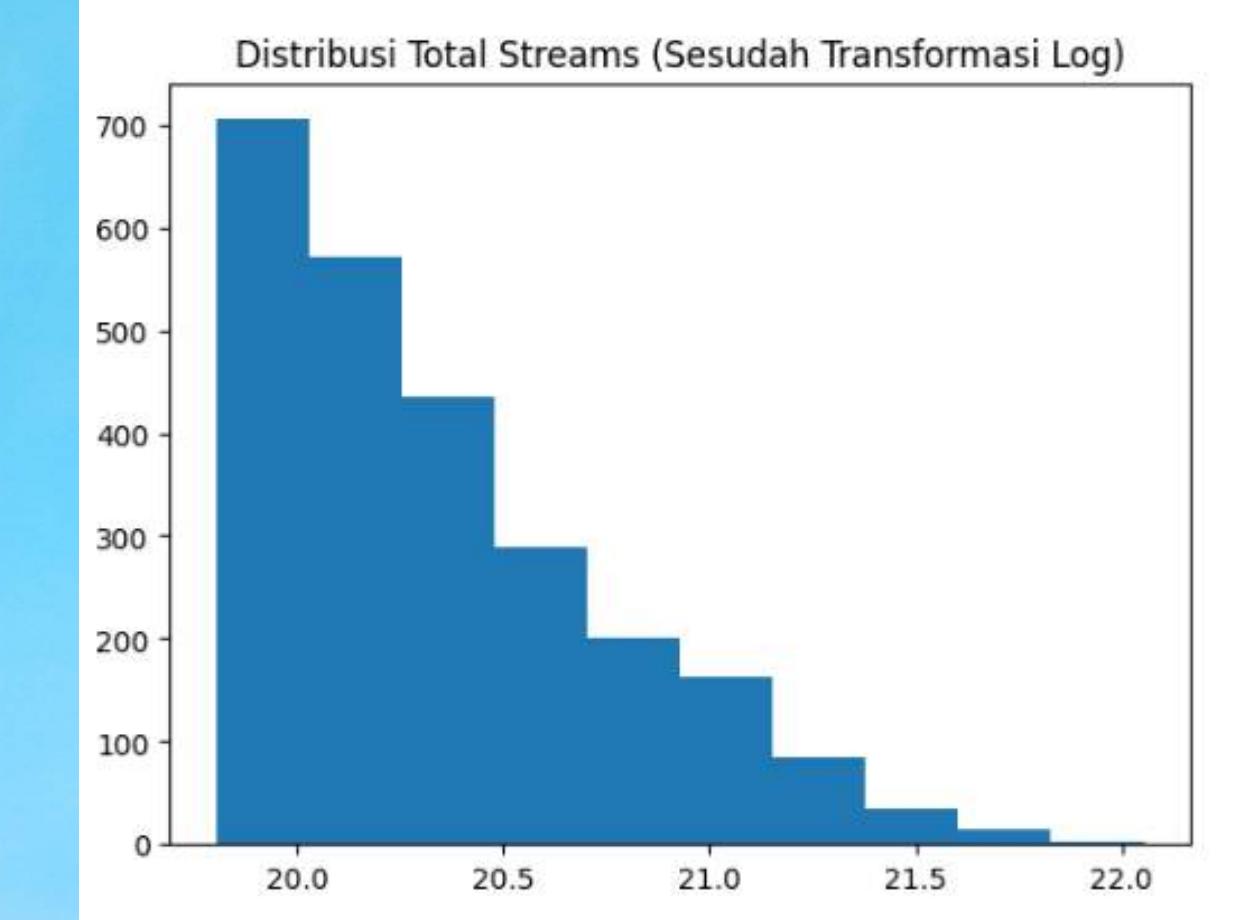


# PREPROCES SINGDATA



## 1. Distribusi Total Streams (Sebelum Transformasi Log)

- Grafik menunjukkan distribusi total streams sangat tidak merata (right-skewed)
- Sebagian besar lagu memiliki jumlah streaming relatif kecil
- Beberapa lagu memiliki nilai streaming sangat besar (outlier)
- Kondisi ini dapat memengaruhi perhitungan jarak pada algoritma K-Means



## 2. Distribusi Total Streams (Sesudah Transformasi Log)

- Setelah dilakukan transformasi logaritmik, distribusi data menjadi lebih seimbang
- Pengaruh nilai ekstrem (outlier) berkurang
- Pola data lebih mendekati distribusi normal

# INTERPRETASI CLUSTER

## Cluster 0 – Lagu Super-Trending

Cluster ini berisi lagu-lagu dengan aktivitas streaming harian sangat tinggi dibandingkan total streams yang dimiliki. Lagu pada cluster ini sedang berada pada puncak popularitas dan memiliki momentum pertumbuhan yang kuat, sehingga dapat dikategorikan sebagai lagu yang sedang sangat diminati pendengar saat ini.

## Cluster 1 – Lagu Lama dengan Popularitas Menurun

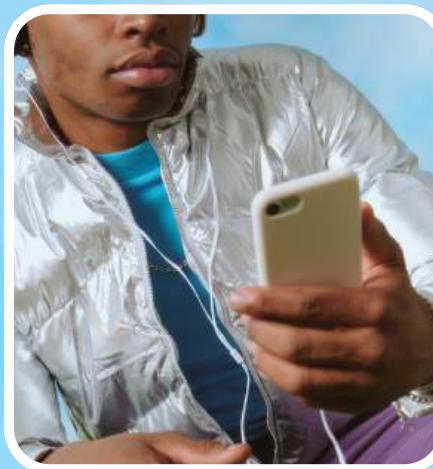
Cluster ini didominasi oleh lagu-lagu yang memiliki total streams cukup besar, namun streams harian relatif rendah. Hal ini menunjukkan bahwa lagu-lagu tersebut pernah populer di masa lalu, tetapi saat ini intensitas pemutarannya mulai menurun.

## Cluster 2 – Mega Hits

Cluster ini mencakup lagu-lagu dengan total streams paling tinggi dan tingkat popularitas kumulatif yang sangat besar. Meskipun aktivitas streaming hariannya tidak setinggi cluster super-trending, lagu-lagu dalam cluster ini tetap memiliki daya tarik kuat dan basis pendengar yang luas.

## Cluster 3 – Lagu Stabil dan Konsisten

Cluster ini berisi lagu-lagu dengan popularitas menengah hingga tinggi serta streams harian yang stabil. Lagu pada cluster ini tidak mengalami lonjakan ekstrem, namun tetap konsisten diputar oleh pendengar dan memiliki performa yang relatif seimbang.



# IMPLIKASI PENELITIAN



## Strategi Promosi Lagu

Hasil clustering dapat digunakan untuk menentukan prioritas promosi, di mana lagu pada cluster super-trending dapat dipromosikan lebih intensif karena memiliki potensi pertumbuhan yang tinggi.



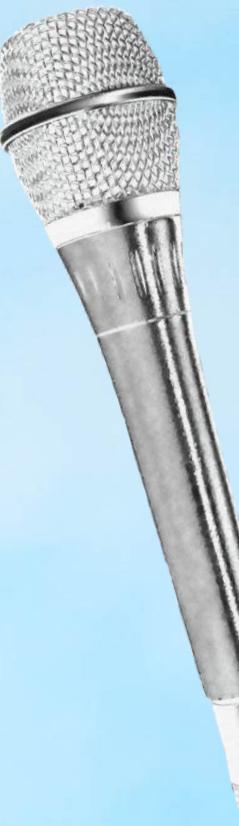
## Kurasi Playlist

Pengelompokan lagu membantu platform streaming dalam menyusun playlist yang lebih tepat, seperti playlist lagu trending, mega hits, atau lagu dengan performa stabil.



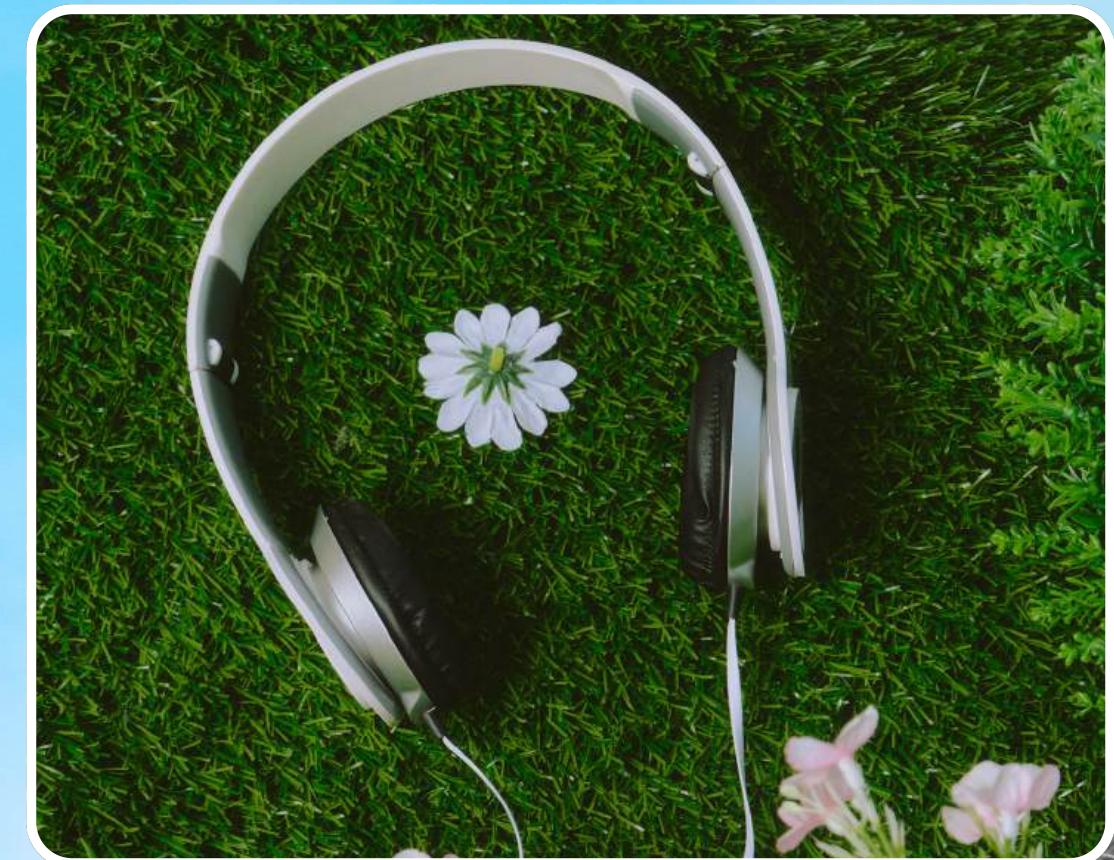
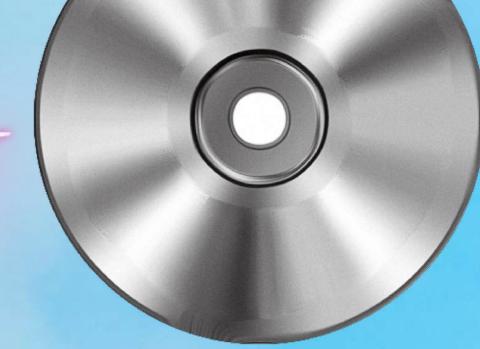
## Pengambilan Keputusan Industri Musik

Label rekaman dan artis dapat menggunakan hasil penelitian ini untuk memahami dinamika popularitas lagu dan menyusun strategi rilis serta distribusi konten secara lebih terarah.





# KETERBATASAN PENELITIAN



- **Keterbatasan Fitur**

- Penelitian ini hanya menggunakan indikator popularitas streaming seperti total streams dan streams harian, sehingga belum mempertimbangkan faktor lain seperti genre, waktu rilis, atau wilayah pendengar.

- **Metode Clustering**

- Algoritma yang digunakan hanya K-Means, yang sensitif terhadap skala data dan outlier serta mengasumsikan bentuk cluster tertentu, sehingga hasil pengelompokan mungkin belum sepenuhnya optimal.

- **Aspek Temporal**

- Analisis dilakukan pada satu periode data, sehingga belum dapat menggambarkan perubahan popularitas lagu dari waktu ke waktu secara dinamis.

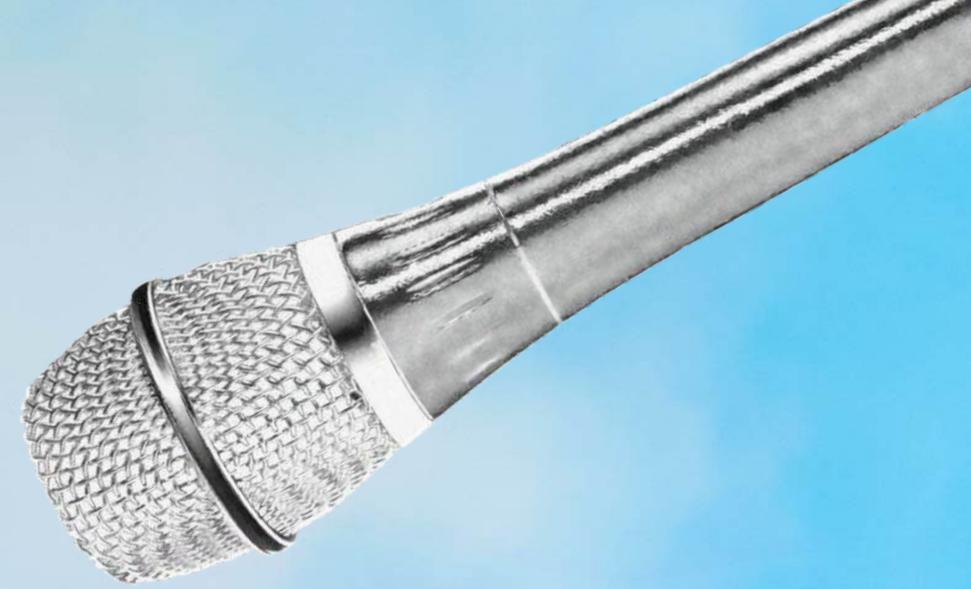
- **Keterbatasan Interpretasi**

- Hasil clustering memberikan gambaran umum pola popularitas lagu, namun tidak secara langsung menjelaskan penyebab meningkat atau menurunnya popularitas suatu lagu.



# KESIMPULAN

Berdasarkan hasil analisis yang telah dilakukan, dapat disimpulkan bahwa algoritma K-Means mampu mengelompokkan lagu-lagu terpopuler Spotify berdasarkan indikator popularitas streaming secara efektif. Melalui tahapan Exploratory Data Analysis (EDA) dan preprocessing, kualitas data berhasil ditingkatkan sehingga proses clustering menghasilkan kelompok lagu yang lebih stabil dan informatif. Pemilihan empat cluster memberikan gambaran yang lebih rinci mengenai pola popularitas lagu, mulai dari lagu yang sedang sangat diminati, lagu dengan popularitas stabil, lagu dengan total pemutaran sangat tinggi, hingga lagu yang popularitasnya mulai menurun. Hasil penelitian ini diharapkan dapat menjadi dasar analisis dan pengambilan keputusan dalam industri musik digital, khususnya dalam mendukung strategi promosi, kurasi playlist, dan pengembangan sistem rekomendasi musik.



# Conclusion

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat.

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat.





# Thank You *For Your Attention*



+123-456-7890



[www.reallygreatsite.com](http://www.reallygreatsite.com)



[hello@reallygreatsite.com](mailto:hello@reallygreatsite.com)

