

ANALISA TINGKAT KEPOPULERAN LAGU BERDASARKAN CLUSTERING ATRIBUT MUSIK

KELOMPOK 5





LATAR BELAKANG

Platform seperti Spotify, Apple Music, dan Shazam menjadi gerbang utama dalam konsumsi musik global. Dataset yang dianalisis mencakup metrik dari ketiga platform ini, termasuk atribut audio (tempo, nada, energi, danceability), popularitas, genre, tanggal rilis, serta data keterlibatan di playlist dan chart.

Analisis ini memberikan wawasan mendalam tentang tren musik dan preferensi pendengar, serta dampak teknologi digital terhadap perilaku konsumsi musik. Temuan ini bermanfaat bagi musisi, label, produser, dan pengembang aplikasi musik dalam memahami dinamika industri secara global.



RUMUSAN MASALAH

- ♪ Bagaimana lagu-lagu di berbagai streaming platform dapat dikelompokkan (di-cluster) menjadi segmen-semen yang berbeda berdasarkan kombinasi atribut-atribut audionya?
- ♪ Bagaimana atribut audio (misalnya, BPM, energi, danceability) berkorelasi dengan popularitas lagu di berbagai streaming platform?
- ♪ Bagaimana keterkaitan antara clustering tipe lagu dengan banyaknya stream di streaming platform?



TUJUAN

- ▶ Menganalisis korelasi antara atribut audio (seperti BPM, energi, dan danceability) dengan tingkat popularitas lagu di platform Spotify.
- ▶ Mengelompokkan (melakukan clustering) lagu-lagu di Spotify ke dalam segmen-semen musik yang berbeda berdasarkan kombinasi atribut-atribut audionya.
- ▶ Menganalisis dan membandingkan rata-rata jumlah streams (sebagai indikator popularitas) di antara segmen-semen musik (cluster) yang berbeda.



METODOLOGI

④ Pengumpulan Data

Dataset yang digunakan yaitu Popular Spotify Songs yang berupa file .csv yang berisi tentang lagu-lagu populer di berbagai platform streaming musik.

Data mencakup lebih dari 20 atribut seperti `track_name`, `artist_name`, `released_year`, `streams`, `bpm`, `danceability_%`, `energy_%`, `valence_%`, `acousticness_%`, `instrumentalness_%`, `liveness_%`, `speechiness_%` dan dimuat menggunakan Python menggunakan library pandas.

⑥ Tools dan Lingkungan Analisis

Proses analisis menggunakan Google Collab. Tools dan libraries yang digunakan meliputi :

- `pandas`
- `matplotlib` dan `seaborn`
- `sklearn`
- `numpy`
- `StandardScaler`

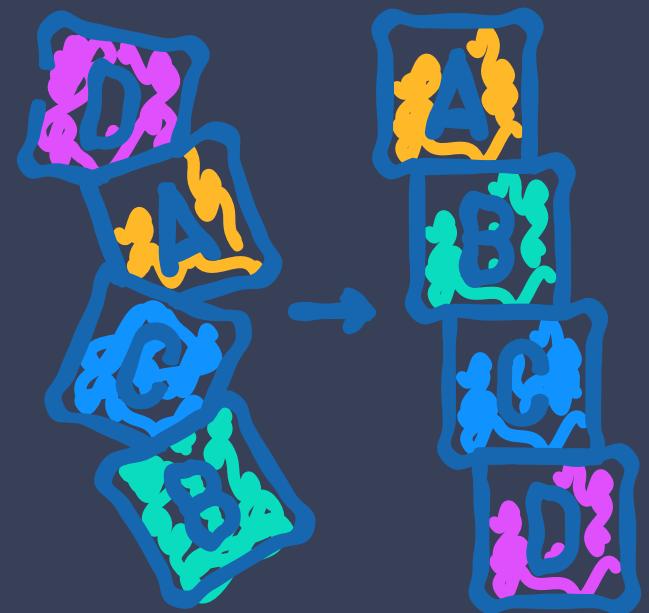


Pembersihan Data dan Eksplorasi Data (EDA)

- Langkah awal yaitu mengecek jumlah data kosong dan duplikat menggunakan `df.isnull().sum()` dan `df.duplicated().sum()`.
- Eksplorasi data dilakukan dengan visualisasi box plot untuk memahami distribusi dan potensi outlier dalam fitur numerik.
- Grafik batang horizontal (barh) digunakan untuk membandingkan lagu-lagu teratas di Spotify dan Apple Music berdasarkan jumlah playlist.
- Visualisasi garis (line chart) juga digunakan untuk menampilkan total stream tahunan dari 5 artis dengan jumlah stream tertinggi,

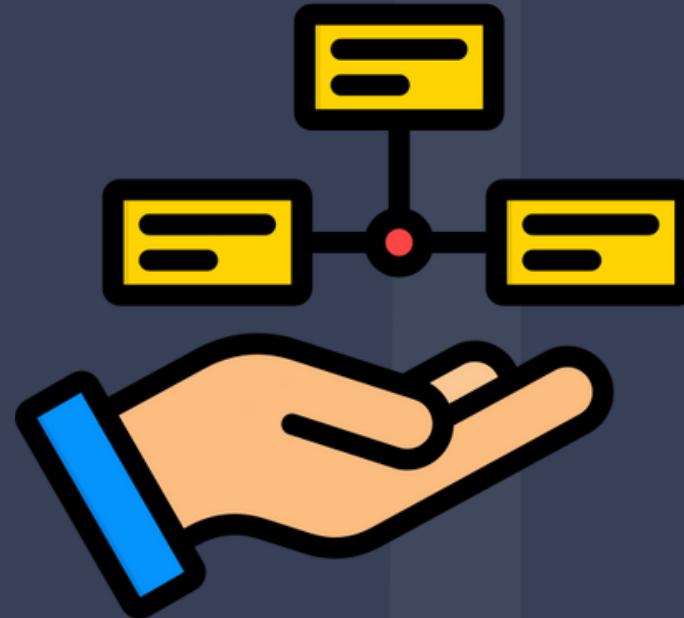


CLUSTERING



Algoritma

K-Means Clustering.



Fitur yang digunakan untuk clustering

danceability, valence, energy, acousticness, instrumentalness, liveness, dan speechiness

Pengambilan Dataset

	track_name	artist(s)_name	artist_count	released_year	released_month	released_day	in_spotify_playlists	in_spotify_charts	streams	in_apple_playlists
0	Seven (feat. Latto) (Explicit Ver.)	Latto, Jung Kook	2	2023	7	14	553	147	141381703	43
1	LALA	Myke Towers	1	2023	3	23	1474	48	133716286	48
2	vampire	Olivia Rodrigo	1	2023	6	30	1397	113	140003974	94
3	Cruel Summer	Taylor Swift	1	2019	8	23	7858	100	800840817	116
4	WHERE SHE GOES	Bad Bunny	1	2023	5	18	3133	50	303236322	84

Dataset yang digunakan merupakan kumpulan lagu populer Spotify dari tahun-tahun terakhir, dengan fitur-fitur seperti nama lagu, nama artis, jumlah artis, tanggal rilis, serta metrik performa seperti jumlah streams, keberadaan di playlist Spotify, Apple Music, dan platform lainnya.

Pengambilan Dataset

	bpm	key	mode	danceability_%	valence_%	energy_%	acousticness_%	instrumentalness_%	liveness_%	speechiness_%
	125	B	Major	80	89	83	31	0	8	4
	92	C#	Major	71	61	74	7	0	10	4
	138	F	Major	51	32	53	17	0	31	6
	170	A	Major	55	58	72	11	0	11	15
	144	A	Minor	65	23	80	14	63	11	6

Pembersihan Dataset

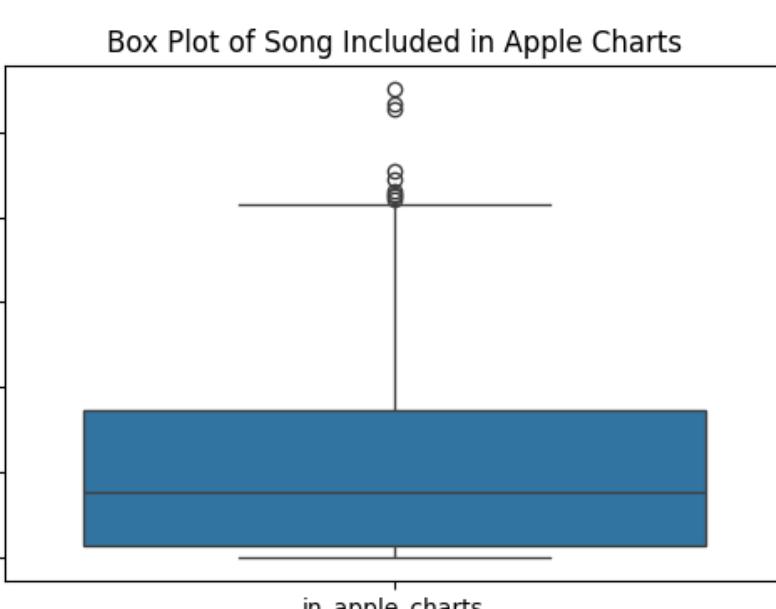
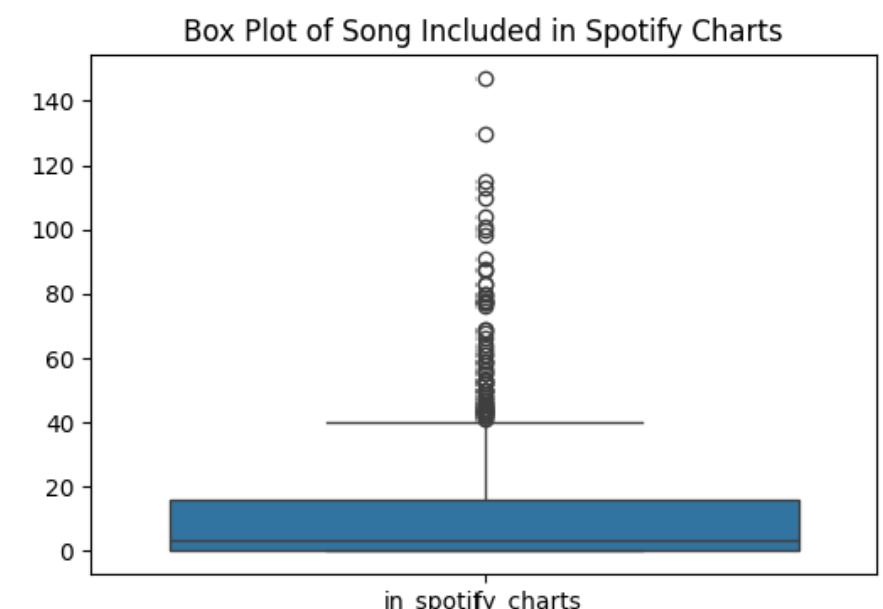
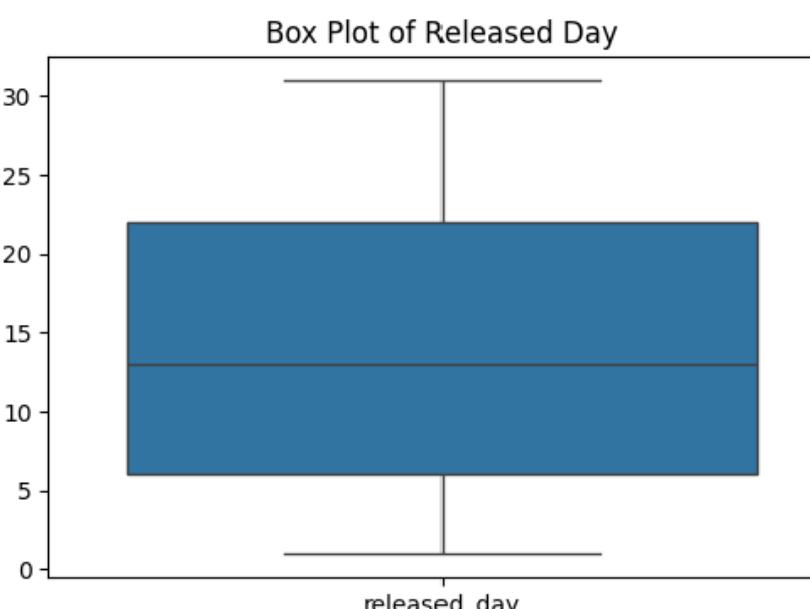
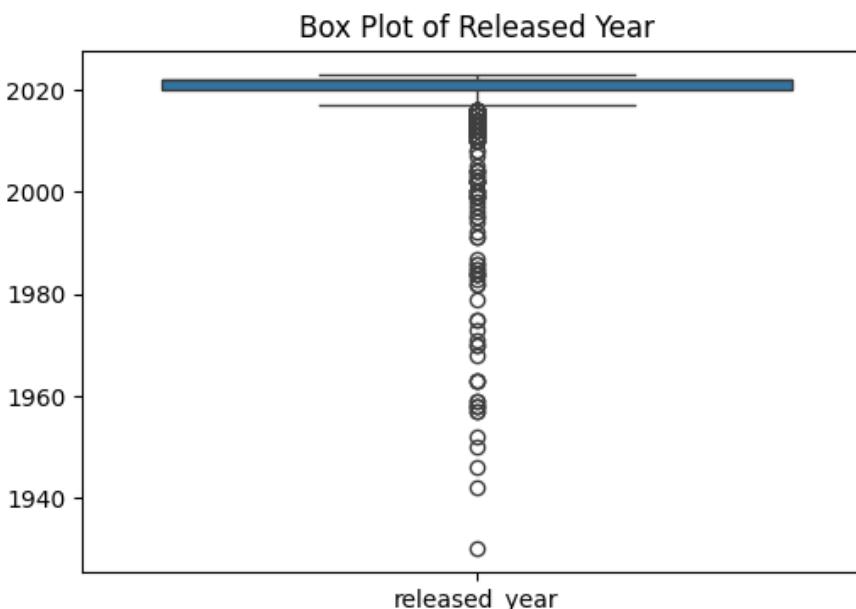
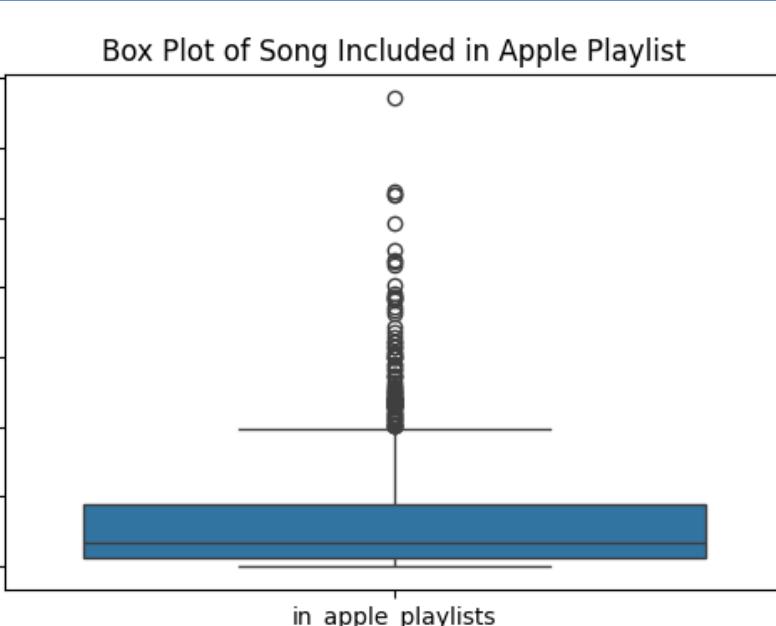
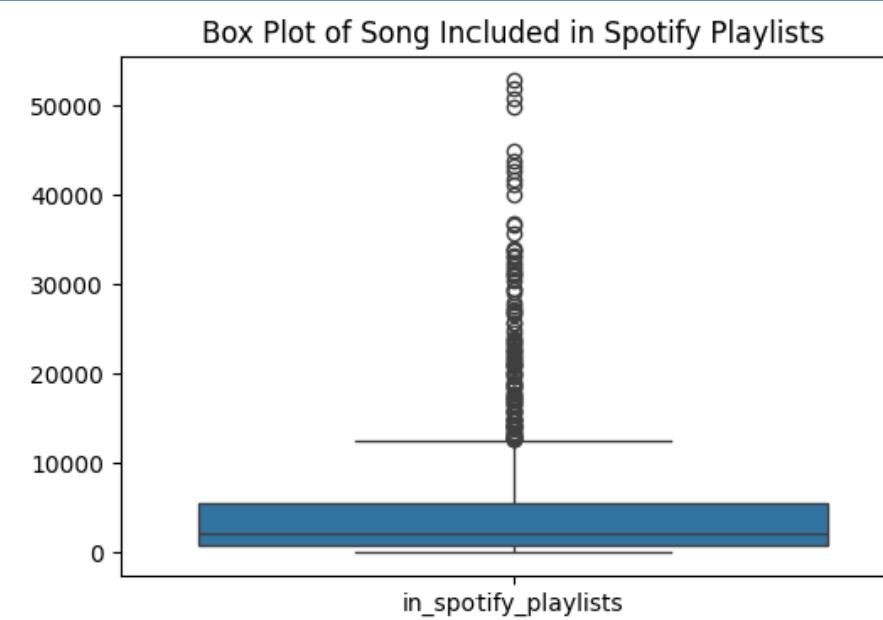
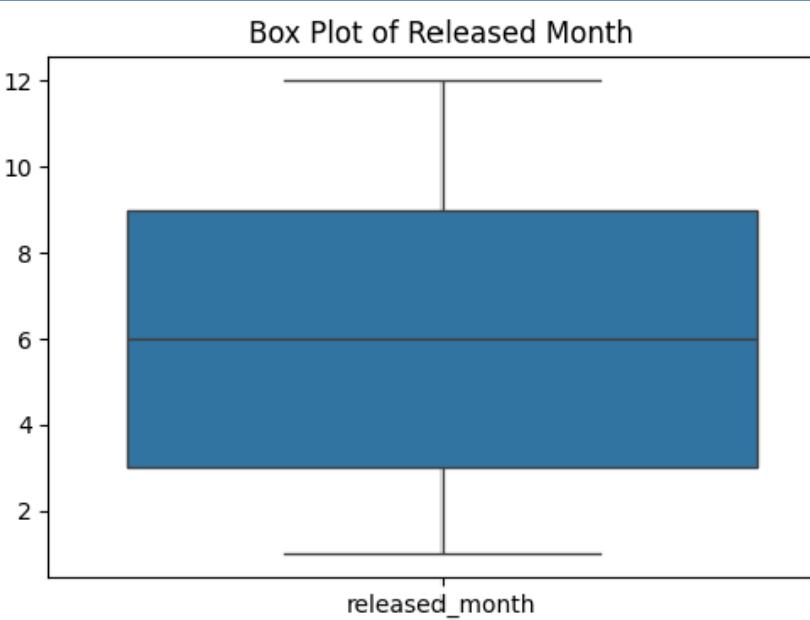
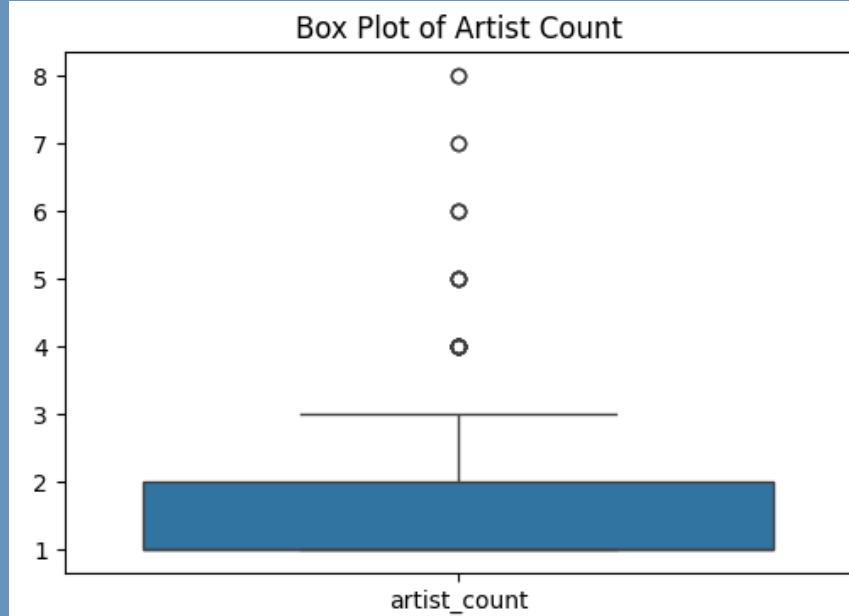
	track_name	artist(s)_name	artist_count	released_year	released_month	released_day	in_spotify_playlists	in_spotify_charts	streams
0	Seven (feat. Latto) (Explicit Ver.)	Latto, Jung Kook	2	2023	7	14	553	147	141381703.0
1	LALA	Myke Towers	1	2023	3	23	1474	48	133716286.0
2	vampire	Olivia Rodrigo	1	2023	6	30	1397	113	140003974.0
3	Cruel Summer	Taylor Swift	1	2019	8	23	7858	100	800840817.0
4	WHERE SHE GOES	Bad Bunny	1	2023	5	18	3133	50	303236322.0

Pembersihan dataset dilakukan dengan mengecek nilai kosong dan duplikat. Beberapa kolom yang tidak relevan, seperti kolom key, dihapus agar analisis lebih fokus. Kolom streams juga dikonversi menjadi tipe numerik agar dapat diproses dengan benar.

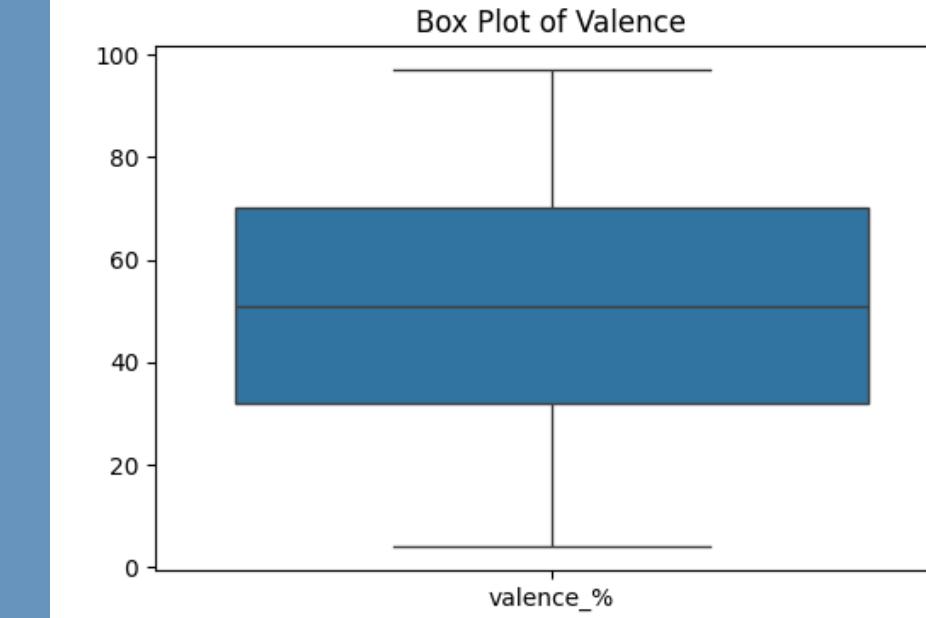
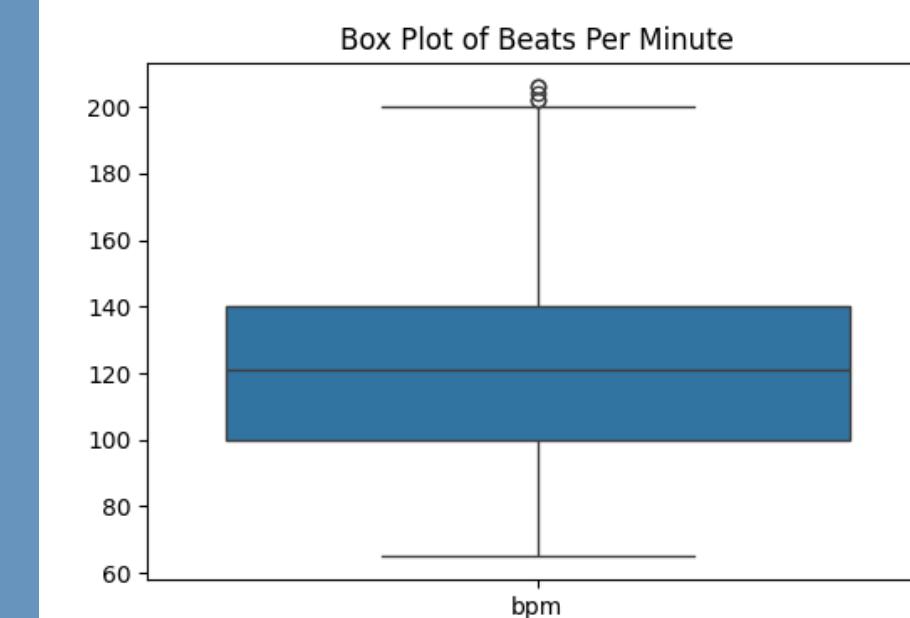
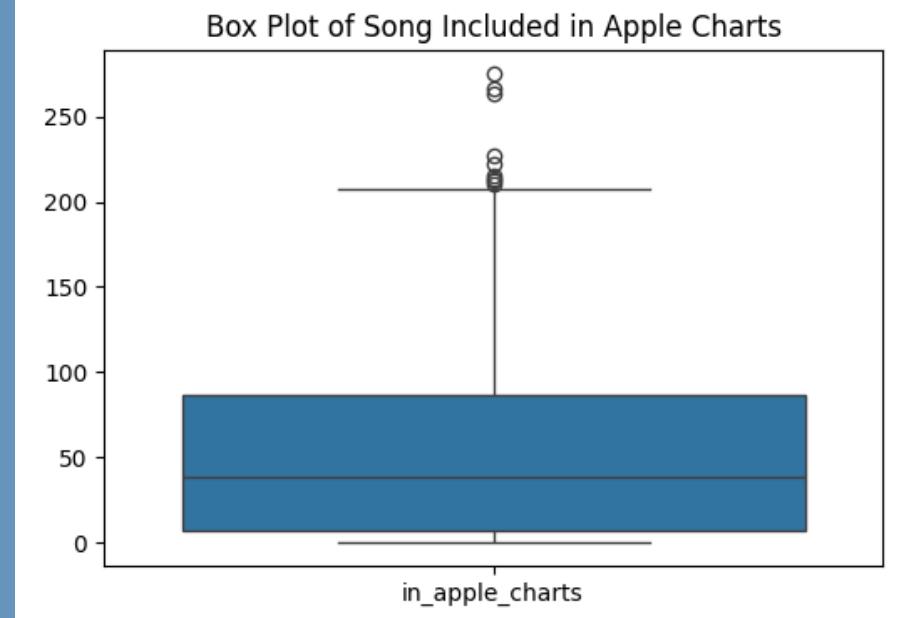
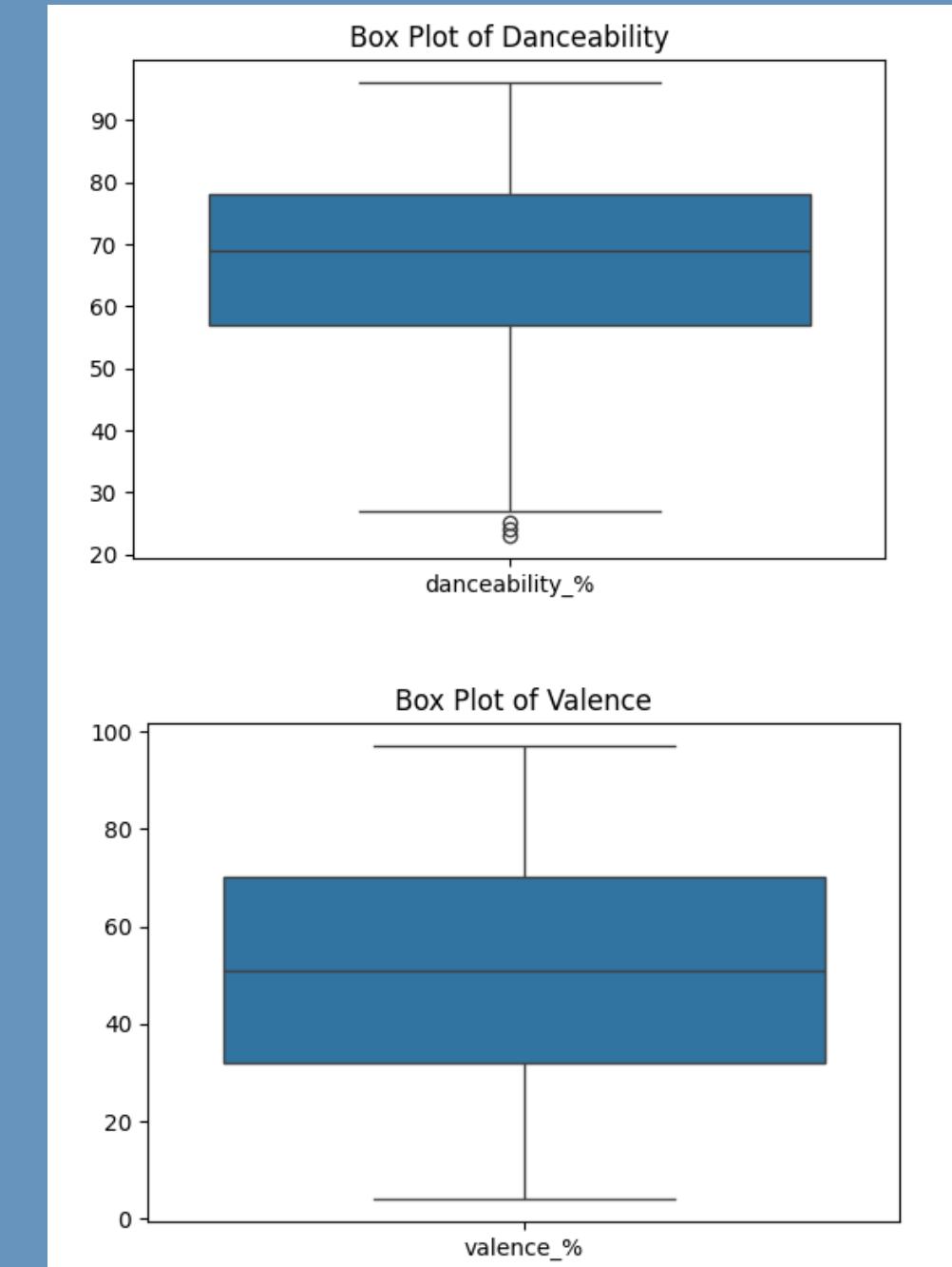
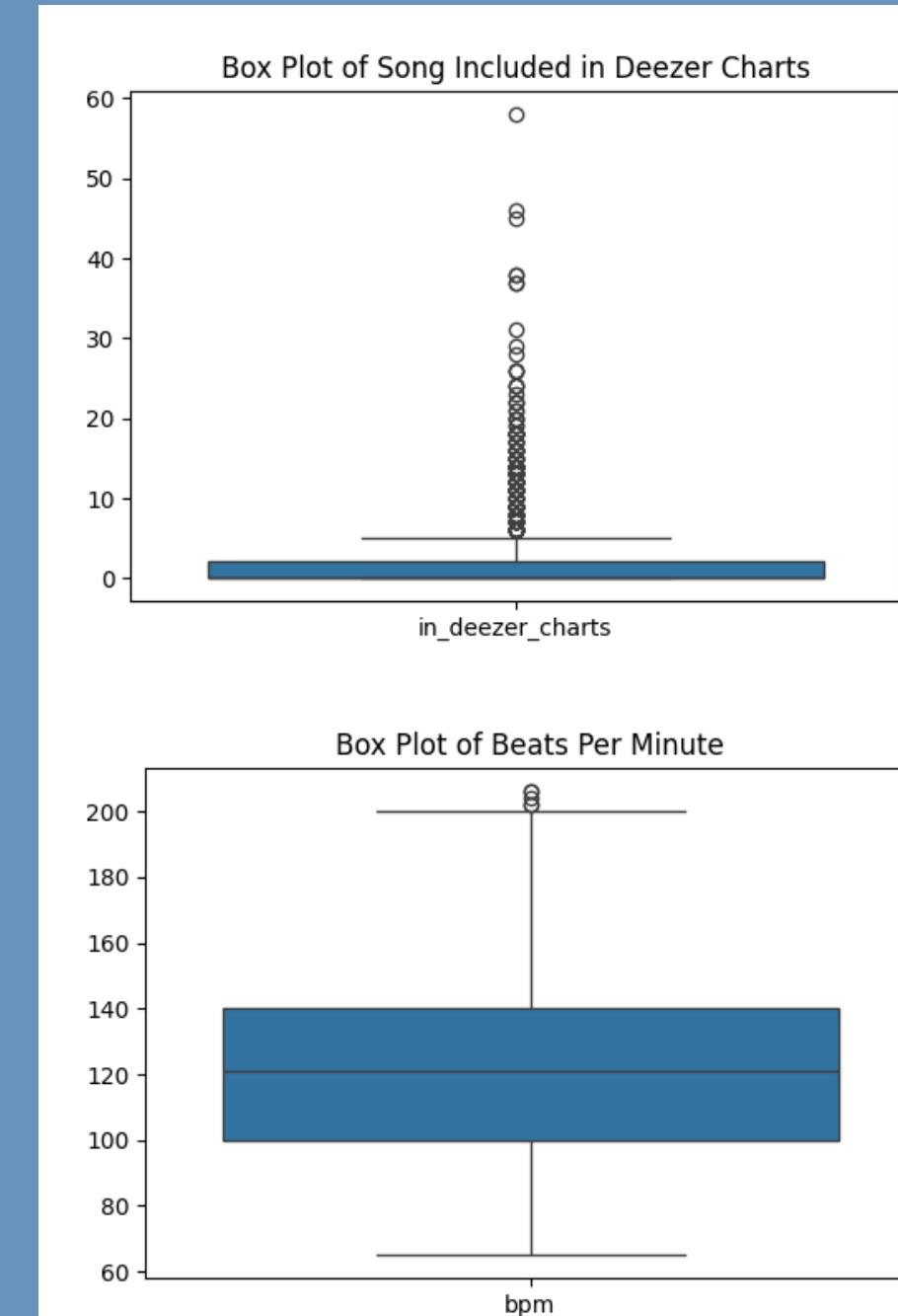
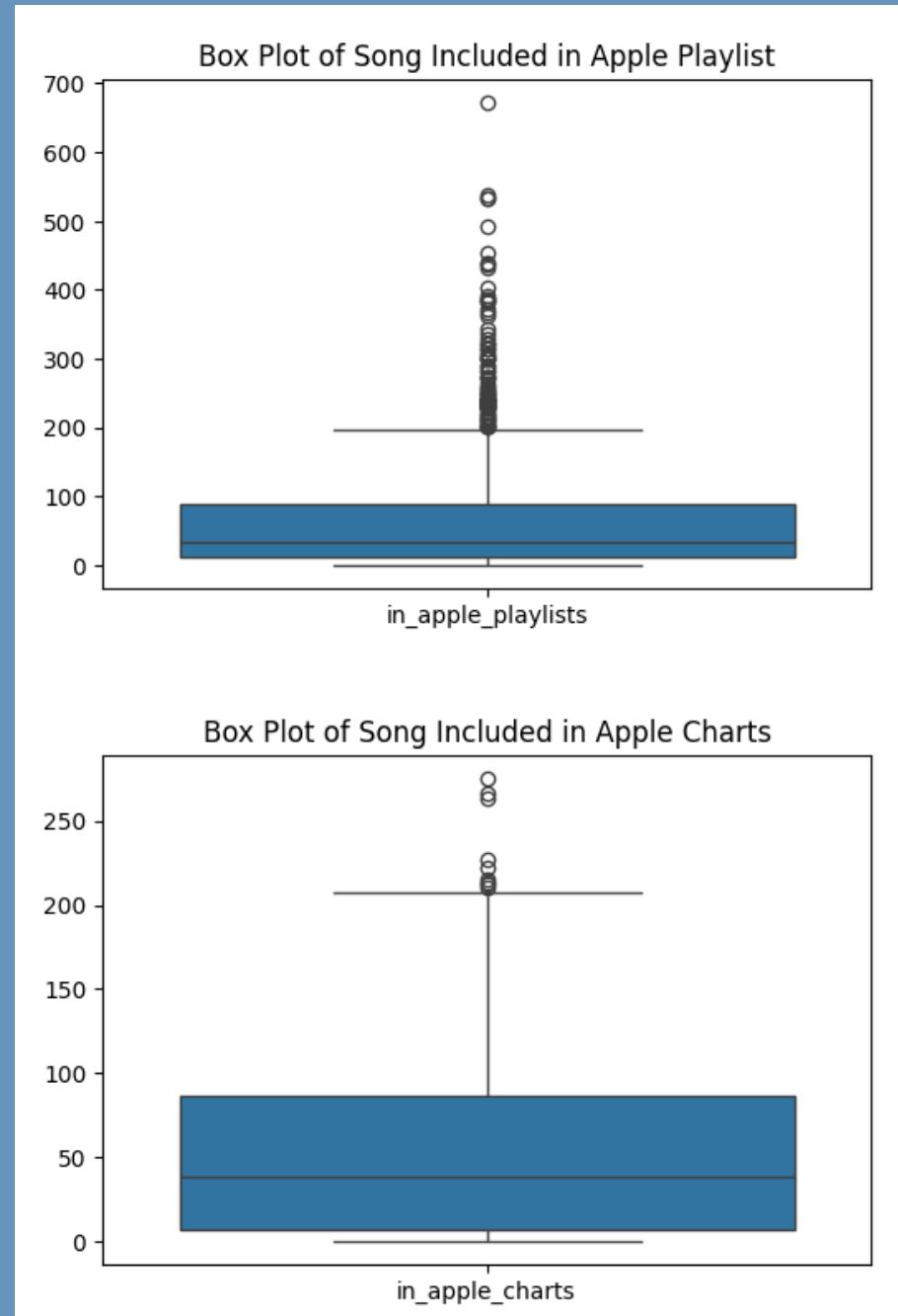
Pembersihan Dataset

in_apple_playlists	...	bpm	mode	danceability_%	valence_%	energy_%	acousticness_%	instrumentalness_%	liveness_%	speechiness_%	Music_Cluster
43	—	125	Major	80	89	83	31	0	8	4	0
48	—	92	Major	71	61	74	7	0	10	4	0
94	—	138	Major	51	32	53	17	0	31	6	2
116	—	170	Major	55	58	72	11	0	11	15	2
84	—	144	Minor	65	23	80	14	63	11	6	4

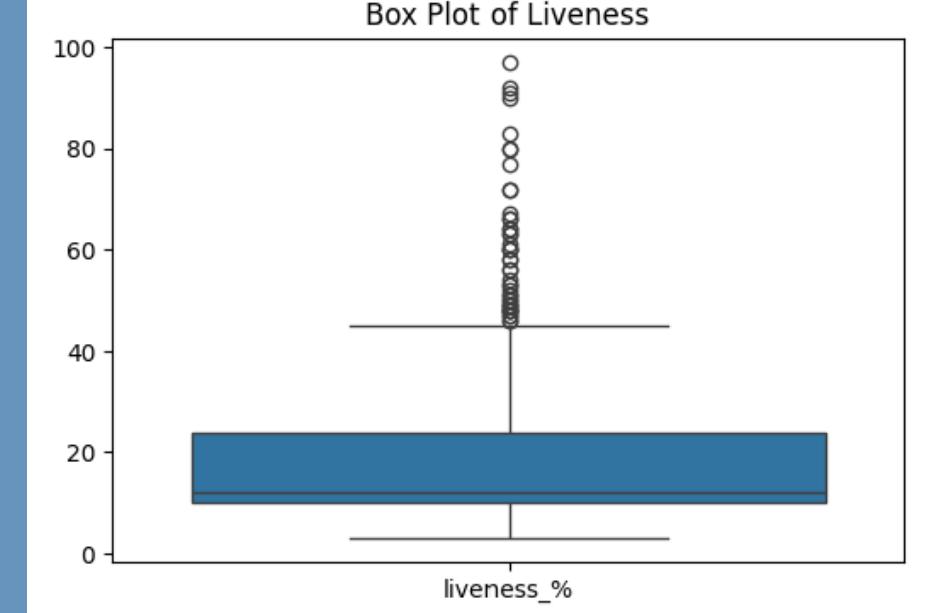
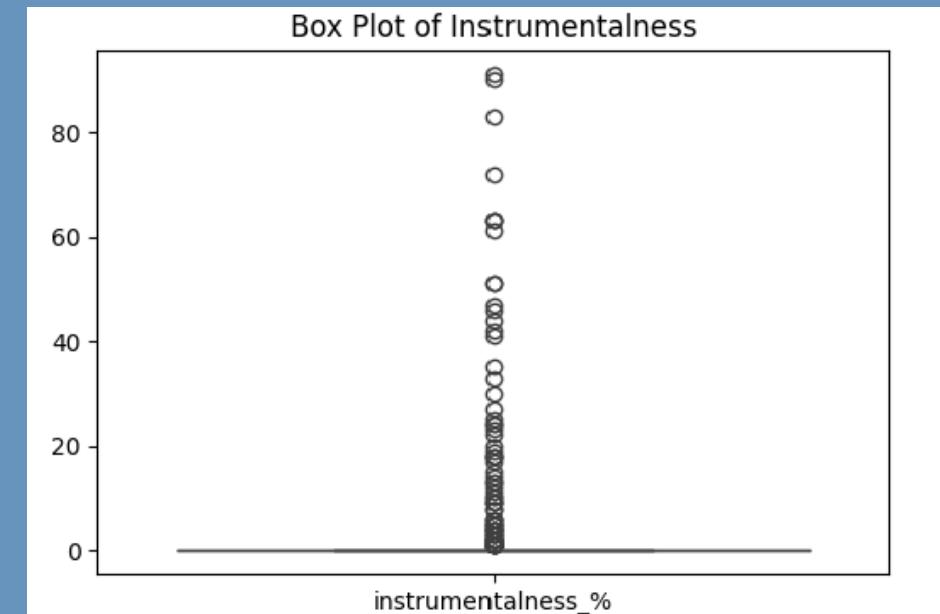
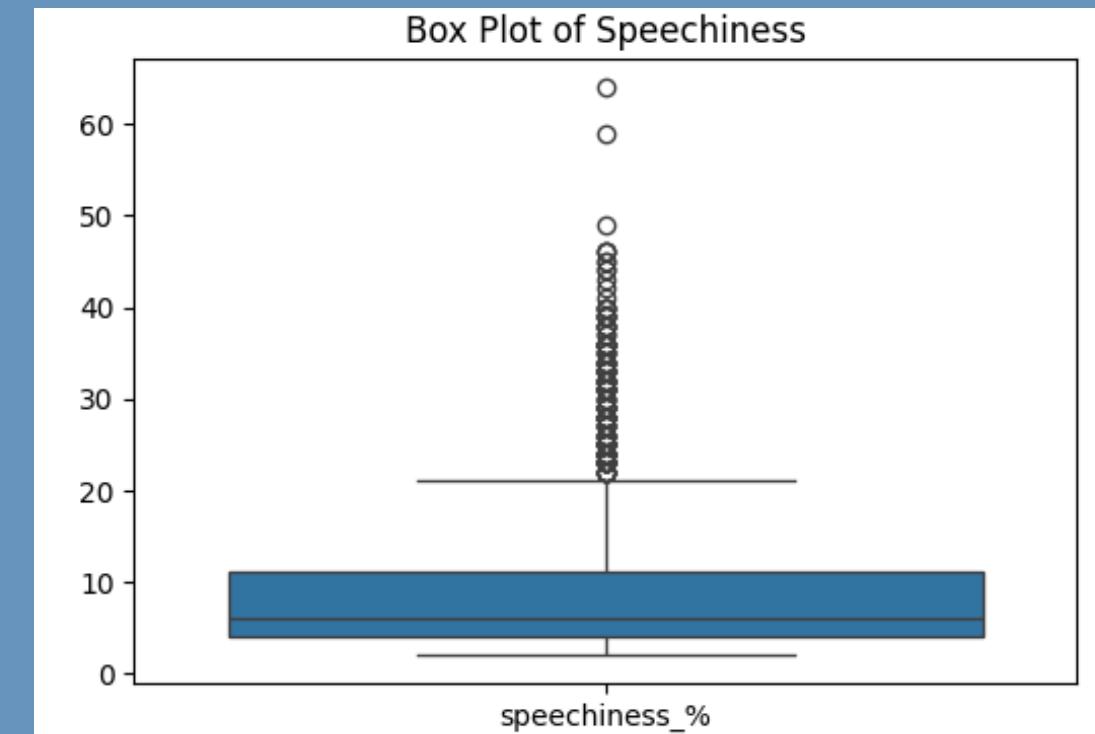
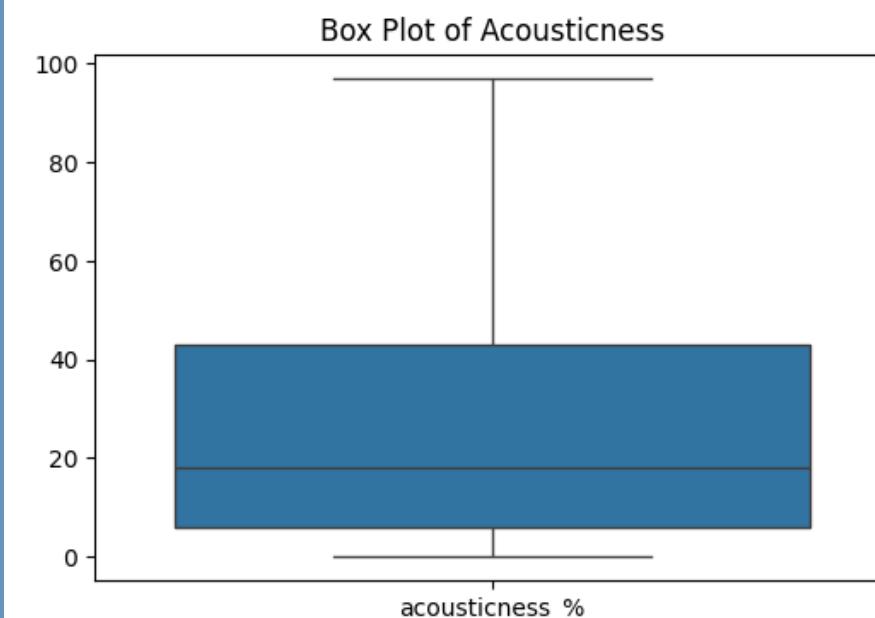
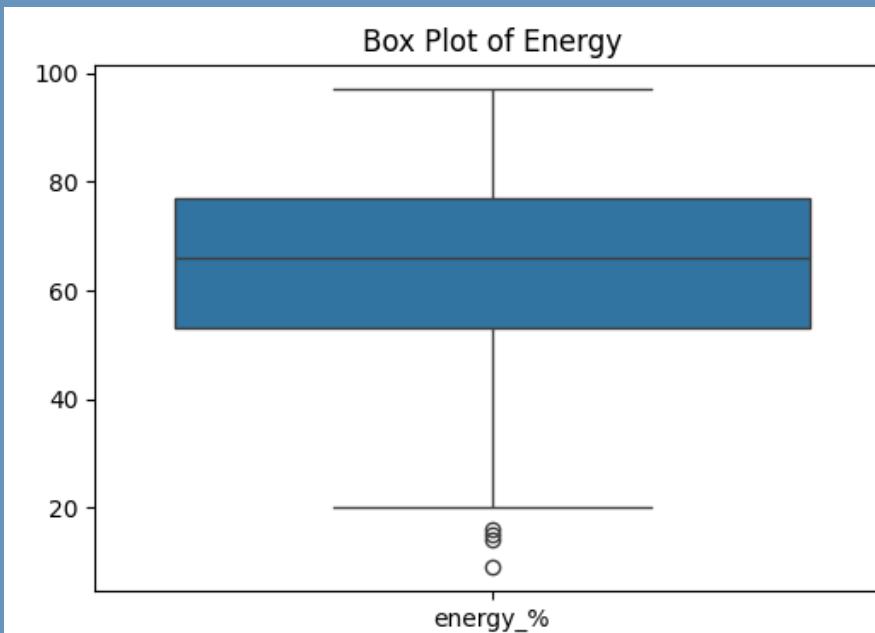
DISTRIBUSI DATA DAN BOXPLOT



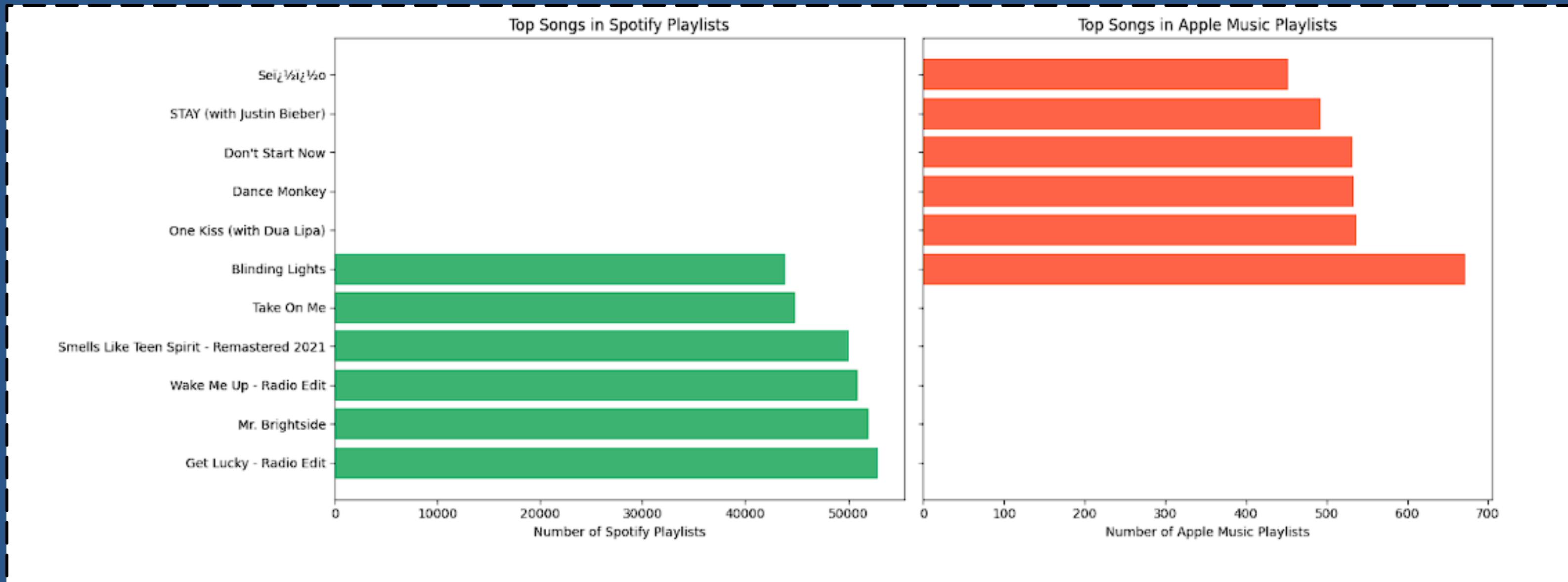
DISTRIBUSI DATA DAN BOXPLOT



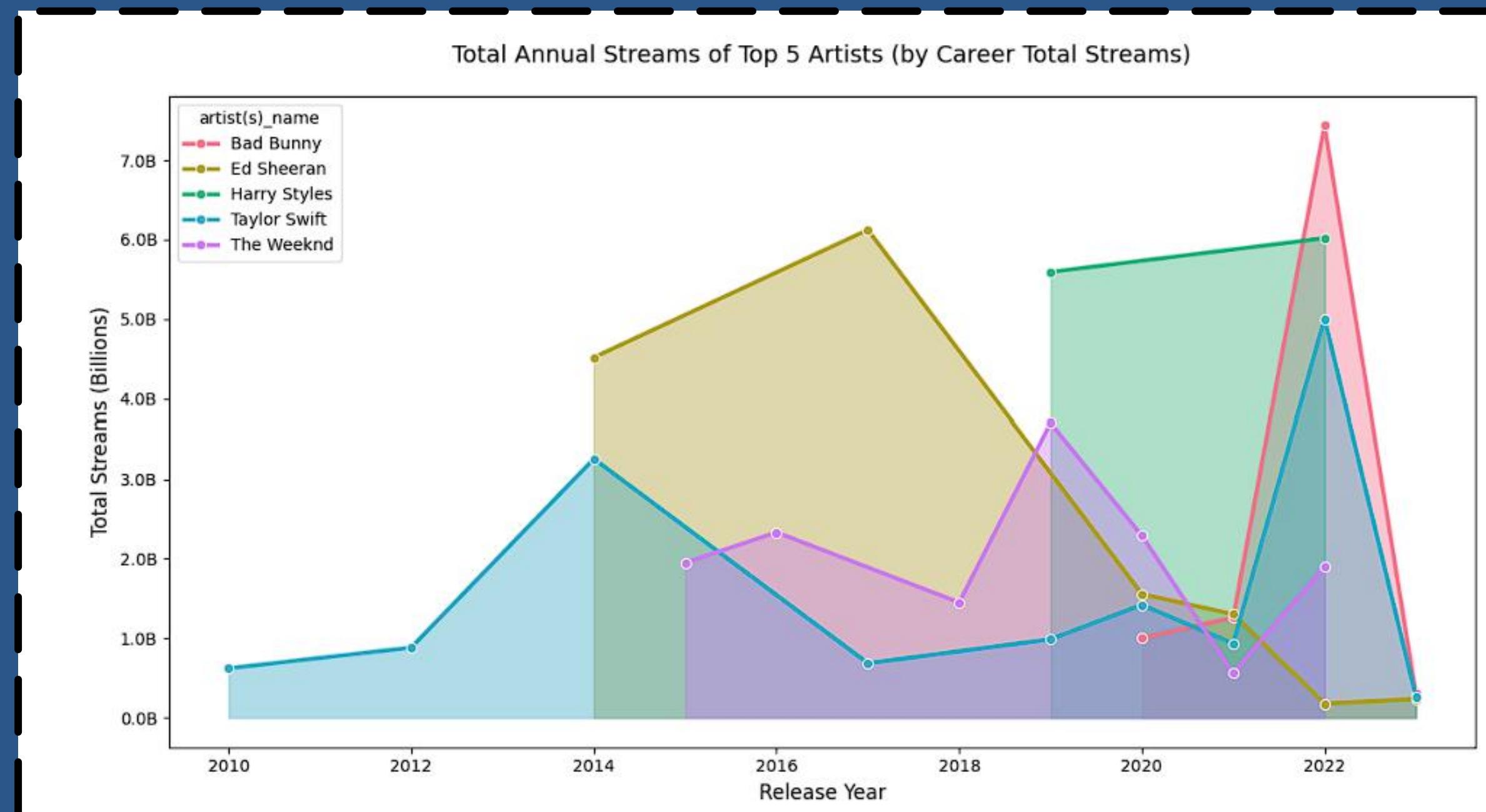
DISTRIBUSI DATA DAN BOXPLOT



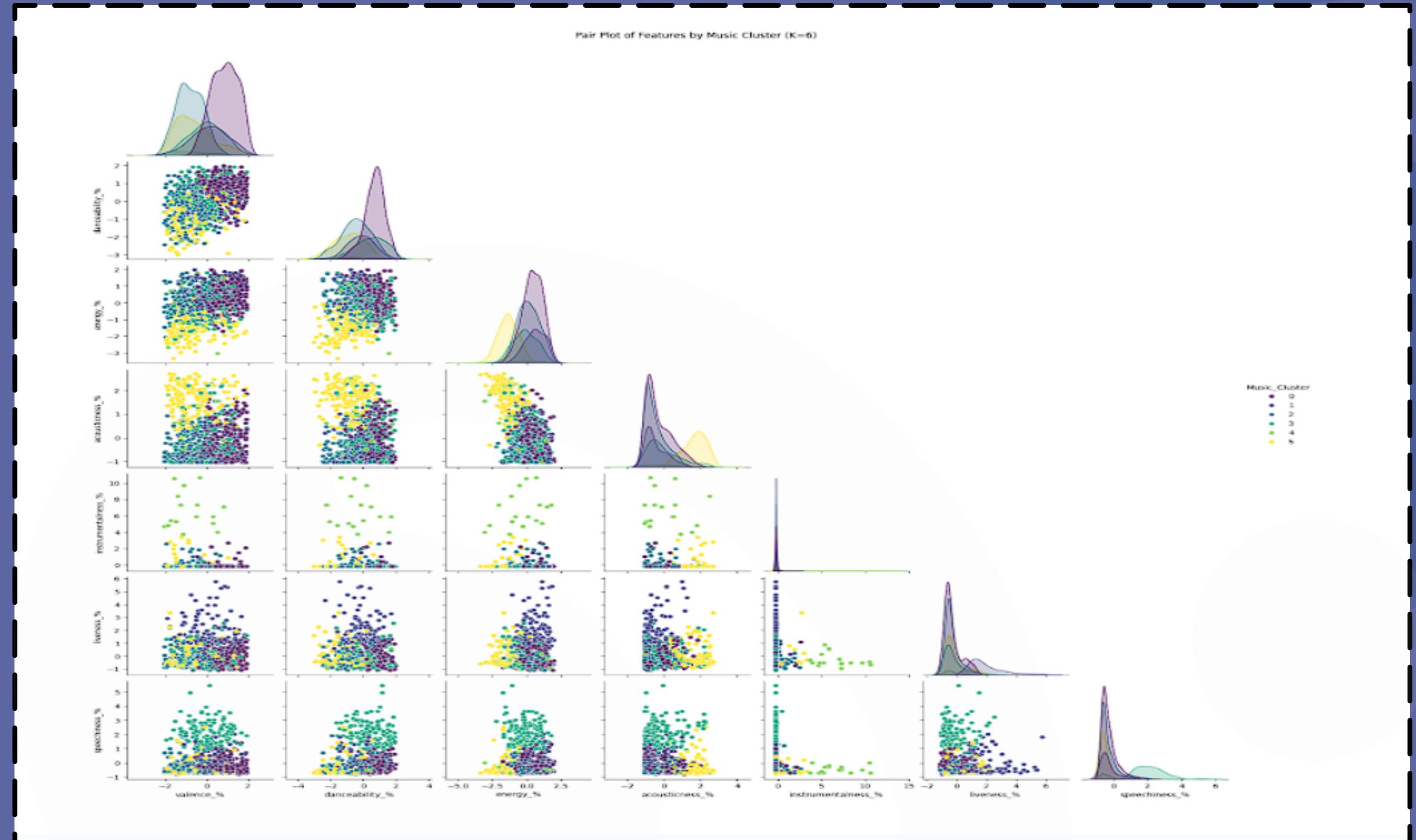
VISUALISASI POPULARITAS LAGU DAN ARTIS



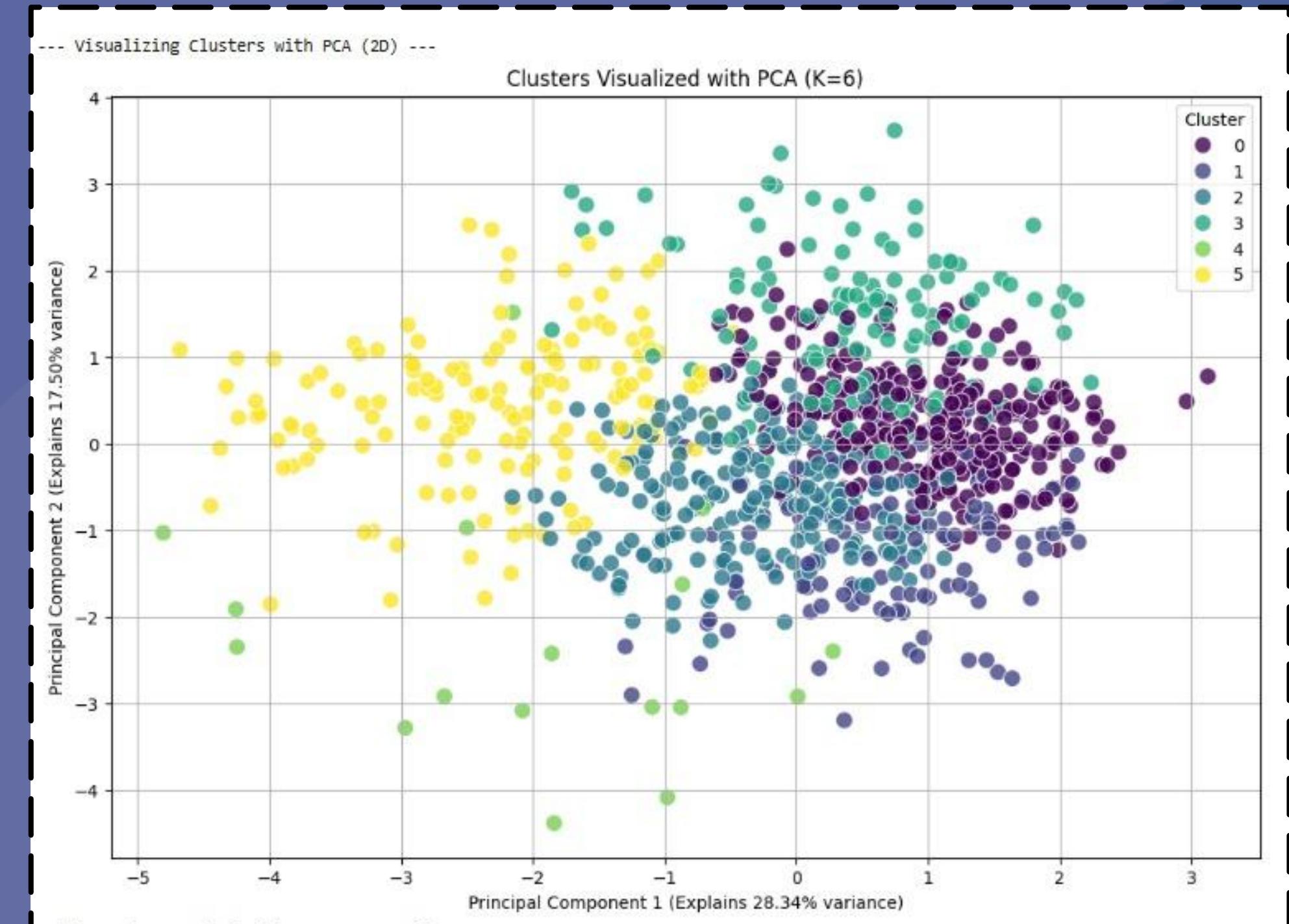
VISUALISASI POPULARITAS LAGU DAN ARTIS



CLUSTERING LAGU DENGAN K-MEANS



CLUSTERING LAGU DENGAN PCA



ANALISIS POPULARITAS SETIAP CLUSTER

Untuk menilai tingkat popularitas dari masing-masing klaster musik, dilakukan analisis rata-rata jumlah streams per klaster. Hasilnya menunjukkan bahwa klaster "Mellow & Reflective Acoustic Pieces" memiliki rata-rata streams tertinggi, disusul oleh "Intense & Brooding Rhythmic Tracks" dan "Upbeat & Joyful Dance Hits". Hal ini mengindikasikan bahwa musik dengan nuansa tenang dan emosional ternyata memiliki daya tarik yang besar di kalangan pendengar.

Average 'streams' per Music Cluster (Whole Numbers):

Music_Cluster

5	581635370
2	539048559
0	533662613
1	495382415
3	382790853
4	344317106

Name: streams, dtype: int64

Average streams with Cluster Names (Whole Numbers):

Music_Cluster	Cluster_Name	streams
0	Mellow & Reflective Acoustic Pieces	581635370
1	Intense & Brooding Rhythmic Tracks	539048559
2	Upbeat & Joyful Dance Hits	533662613
3	High-Energy Live Performances	495382415
4	Rhythmic & Speech-Driven Grooves	382790853
5	Atmospheric & Melancholic Instrumentals	344317106

ANALISIS POPULARITAS SETIAP CLUSTER

--- Top 5 Songs in Each Cluster (by Streams) ---

--- Upbeat & Joyful Dance Hits (Cluster 0) ---

	track_name	streams
0	Shape of You	3562543890
1	Dance Monkey	2864791672
2	Sunflower - Spider-Man: Into the Spider-Verse	2808096550
3	Believer	2594040133
4	Closer	2591224264

--- High-Energy Live Performances (Cluster 1) ---

	track_name	streams
0	One Dance	2713922350
1	As It Was	2513188493
2	Watermelon Sugar	2322580122
3	MONTERO (Call Me By Your Name)	1735441776
4	In The End	1624165576

--- Intense & Brooding Rhythmic Tracks (Cluster 2) ---

	track_name	streams
0	Blinding Lights	3703895074
1	STAY (with Justin Bieber)	2665343922
2	Perfect	2559529074
3	Sweater Weather	2282771485
4	Something Just Like This	2204080728

4 Something Just Like This 2204080728

--- Rhythmic & Speech-Driven Grooves (Cluster 3) ---

	track_name	streams
0	Starboy	2565529693
1	Love Yourself	2123309722
2	Lose Yourself - Soundtrack Version	1829992958
3	No Role Modelz	1791000570
4	LA CANCII <i>X</i>	1435127549

--- Atmospheric & Melancholic Instrumentals (Cluster 4) ---

	track_name	streams
0	Apocalypse	841749534
1	Freaks	824420218
2	Master of Puppets (Remastered)	704171068
3	After Dark	646886885
4	Alien Blues	370068639

--- Mellow & Reflective Acoustic Pieces (Cluster 5) ---

	track_name	streams
0	Someone You Loved	2887241814
1	Say You Won't Let Go	2420461338
2	lovely - Bonus Track	2355719893
3	Lucid Dreams	2288695111
4	Photograph	2236667932

Selanjutnya, ditampilkan lima lagu teratas dalam setiap klaster berdasarkan jumlah streams-nya. Visualisasi ini memperlihatkan lagu-lagu populer yang menjadi representasi dominan di masing-masing klaster, seperti "Shape of You", "Blinding Lights", dan "Someone You Loved". Daftar ini membantu memahami karakteristik musik dalam setiap klaster.

KESIMPULAN

Melalui analisis clustering menggunakan algoritma K-Means, lagu-lagu populer berhasil dikelompokkan ke dalam enam klaster berdasarkan atribut audio seperti energy, valence, danceability, dan lainnya. Masing-masing klaster menunjukkan karakteristik musical yang unik, seperti lagu-lagu akustik yang mellow, atau lagu-lagu dengan ritme intens dan enerjik. Analisis rata-rata jumlah streams menunjukkan bahwa klaster "Mellow & Reflective Acoustic Pieces" menjadi yang paling populer di antara lainnya. Temuan ini membuktikan bahwa metode unsupervised learning seperti K-Means dapat dimanfaatkan untuk memahami segmentasi musik secara efektif, serta memberikan wawasan tentang preferensi pendengar dalam skala besar.

THANK YOU

