



Spotify

What Makes a Song Popular?

Ching Chen (Timothy) Liu



Project Motivation & Objective



Objective:

This project analyzes how different audio features (danceability, energy, explicit content, etc.) influence a song's popularity and explores whether a song's success is driven by composition or external factors.

Understanding Song Popularity Matters

- I wanted to gain deeper insights into what makes a song popular—is it the music itself, or do other factors play a role?
- Artists, music labels, and streaming platforms need to understand popularity trends to make better decisions in the music business.

Dataset Overview

Comedy	73	230666	FALSE	0.676	0.461	1	-6.746	0	0.143	0.	0	0
Ghost - Acoustic	55	149610	FALSE	0.42	0.166	1	-17.235	1	0.0763	0	0	0
To Begin Again	57	210826	FALSE	0.438	0.359	0	-9.734	1	0.0557	0	0	0
Can't Help Falling In Love	71	201933	FALSE	0.266	0.0596	0	-18.515	1	0.0363	0	0	0
Hold On	82	198853	FALSE	0.618	0.443	2	-9.681	1	0.0526	0	0	0
Days I Will Remember	58	214240	FALSE	0.688	0.481	6	-8.807	1	0.105	0	0	0
Say Something	74	229400	FALSE	0.407	0.147	2	-8.822	1	0.0355	0	0	0
I'm Yours	80	242946	FALSE	0.703	0.444	11	-9.331	1	0.0417	0	0	0
Lucky	74	189613	FALSE	0.625	0.414	0	-8.7	1	0.0369	0	0	0
Hunger	56	205594	FALSE	0.442	0.632	1	-6.77	1	0.0295	0	0	0
Give Me Your Forever	74	244800	FALSE	0.627	0.363	8	-8.127	1	0.0291	0	0	0
I Won't Give Up	69	240165	FALSE	0.483	0.303	4	-10.058	1	0.0429	0	0	0
Solo	52	198712	FALSE	0.489	0.314	7	-9.245	0	0.0331	0	0	0
Bad Liar	62	248448	FALSE	0.691	0.234	3	-6.441	1	0.0285	0	0	0
Hold On - Remix	56	188133	FALSE	0.755	0.78	2	-6.084	1	0.0327	0	0	0
Falling in Love at a Coffee Shop	58	244986	FALSE	0.489	0.561	4	-7.933	1	0.0274	0	0	0
ily (i love you baby)	56	129750	FALSE	0.706	0.112	2	-18.098	1	0.0391	0	0	0
At My Worst	54	169728	FALSE	0.795	0.0841	10	-18.09	0	0.0461	0	0	0
Lucky	68	189613	FALSE	0.625	0.414	0	-8.7	1	0.0369	0	0	0
Photograph	67	260186	FALSE	0.717	0.32	3	-8.393	1	0.0283	0	0	0
I'm Yours	75	242946	FALSE	0.703	0.444	11	-9.331	1	0.0417	0	0	0
Demons	63	174174	FALSE	0.678	0.351	0	-8.654	1	0.0266	0	0	0
Say Something	70	229400	FALSE	0.407	0.147	2	-8.822	1	0.0355	0	0	0
93 Million Miles	0	216386	FALSE	0.572	0.454	3	-10.286	1	0.0258	0	0	0
Unlonely	0	231266	FALSE	0.796	0.667	5	-4.831	0	0.0392	0	0	0
Bella Luna	1	302346	FALSE	0.755	0.454	9	-9.609	0	0.0352	0	0	0

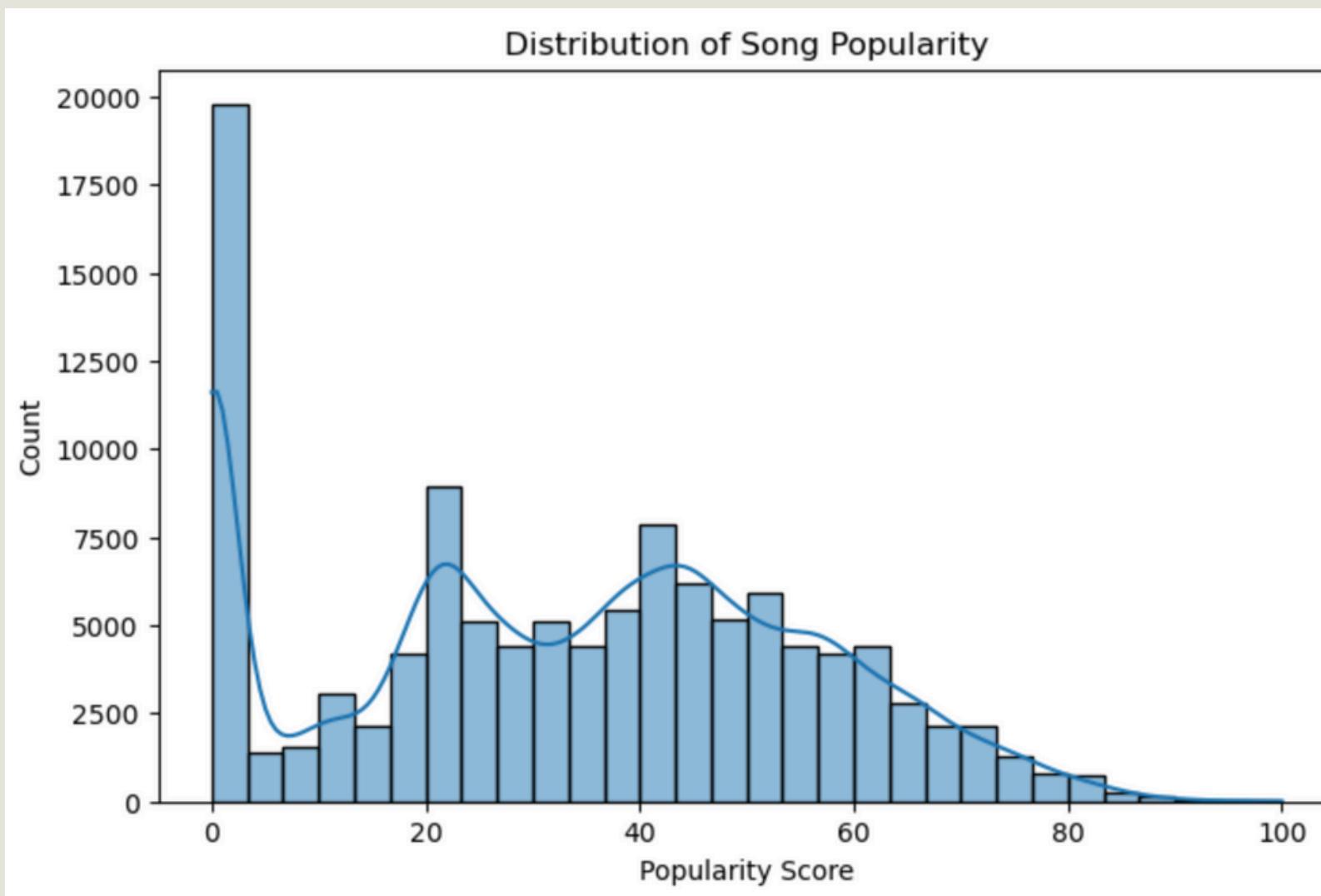
Source

- Spotify dataset
- 114000 Songs

Key Attributes

- Popularity score (0-100) (Target Variable)
- Audio features (danceability, energy, loudness, etc.)
- Explicit content & genre

Popularity Distribution



The histogram below represents the distribution of song popularity scores in our dataset.

Key Observations

- Most songs have very low popularity scores
- A secondary peak appears around the 20-40 popularity range.
- The number of highly popular songs (above 80) is very small.

Business Insights

- Many songs struggle to gain attention, highlighting the importance of marketing & playlist placement.

Correlation Analysis

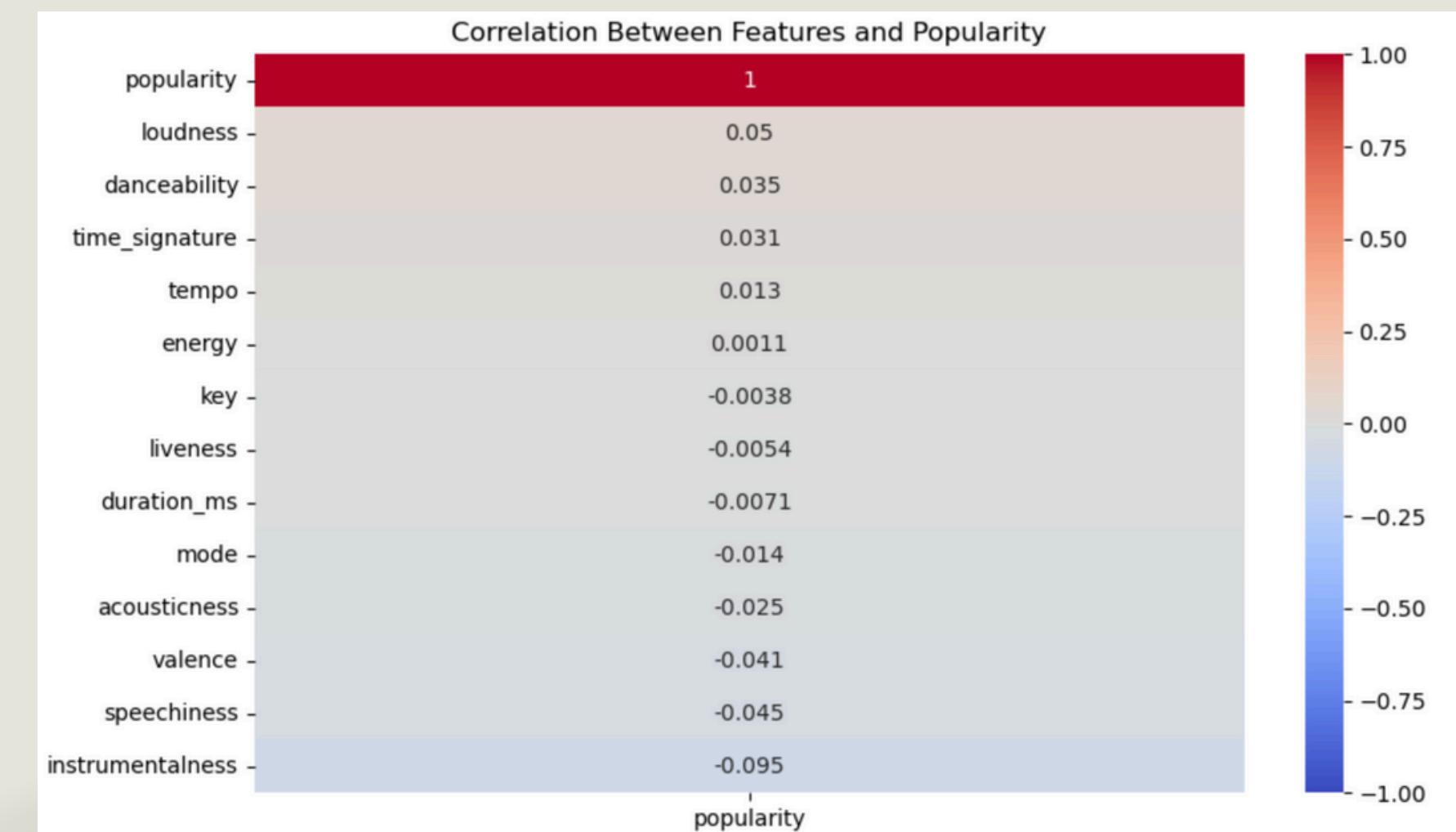
Show heatmap of correlations

Key Takeaways

- No strong correlation between popularity and any single audio feature.
- Loudness (0.05) has the highest positive correlation, suggesting louder songs tend to be slightly more popular.
- The number of highly popular songs (above 80) is very small.

Business Insights

- Popularity is not just about music composition—other factors (e.g., artist fame, promotion) matter more.



Danceability vs. Popularity

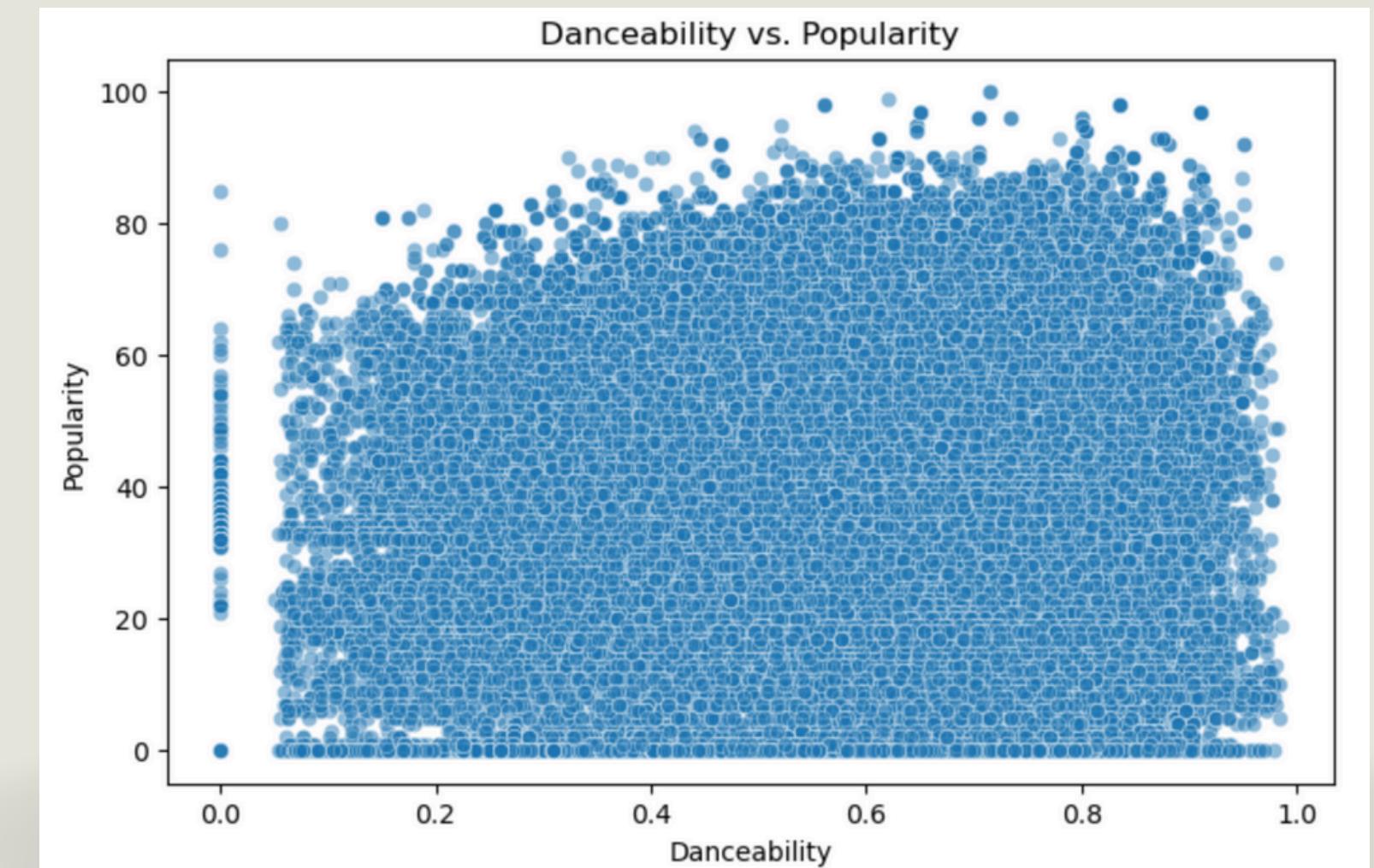
Scatter plot of danceability vs. popularity

Key Takeaways

- There is a weak positive correlation between danceability and popularity.
- Wide popularity range at all danceability levels.
- Many songs have very low popularity regardless of danceability.

Business Insights

- Highly danceable songs might have a better chance at success, but it is not guaranteed.
- Producers & labels should consider marketing, streaming strategy, and artist engagement rather than relying solely on danceability.



Energy vs. Popularity

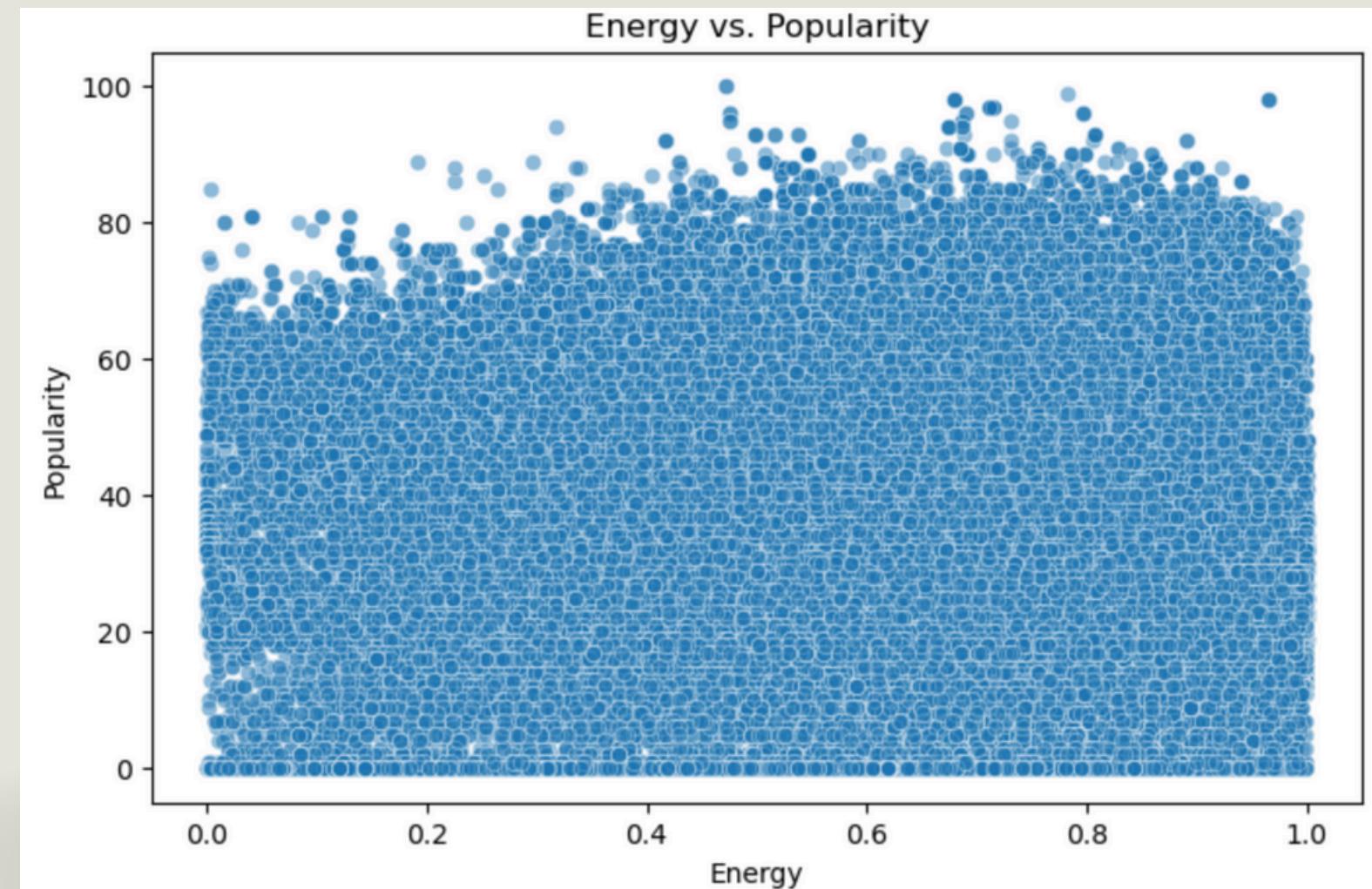
Scatter plot of energy vs. popularity

Key Takeaways

- No strong relationship—low-energy & high-energy songs can both be popular.
- Popularity is likely driven by external factors

Business Insights

- Music producers should not rely on energy alone when crafting hit songs.
- Streaming platforms and record labels should focus more on artist branding, playlist strategy, and listener engagement rather than energy levels.



Explicit Songs vs. Popularity

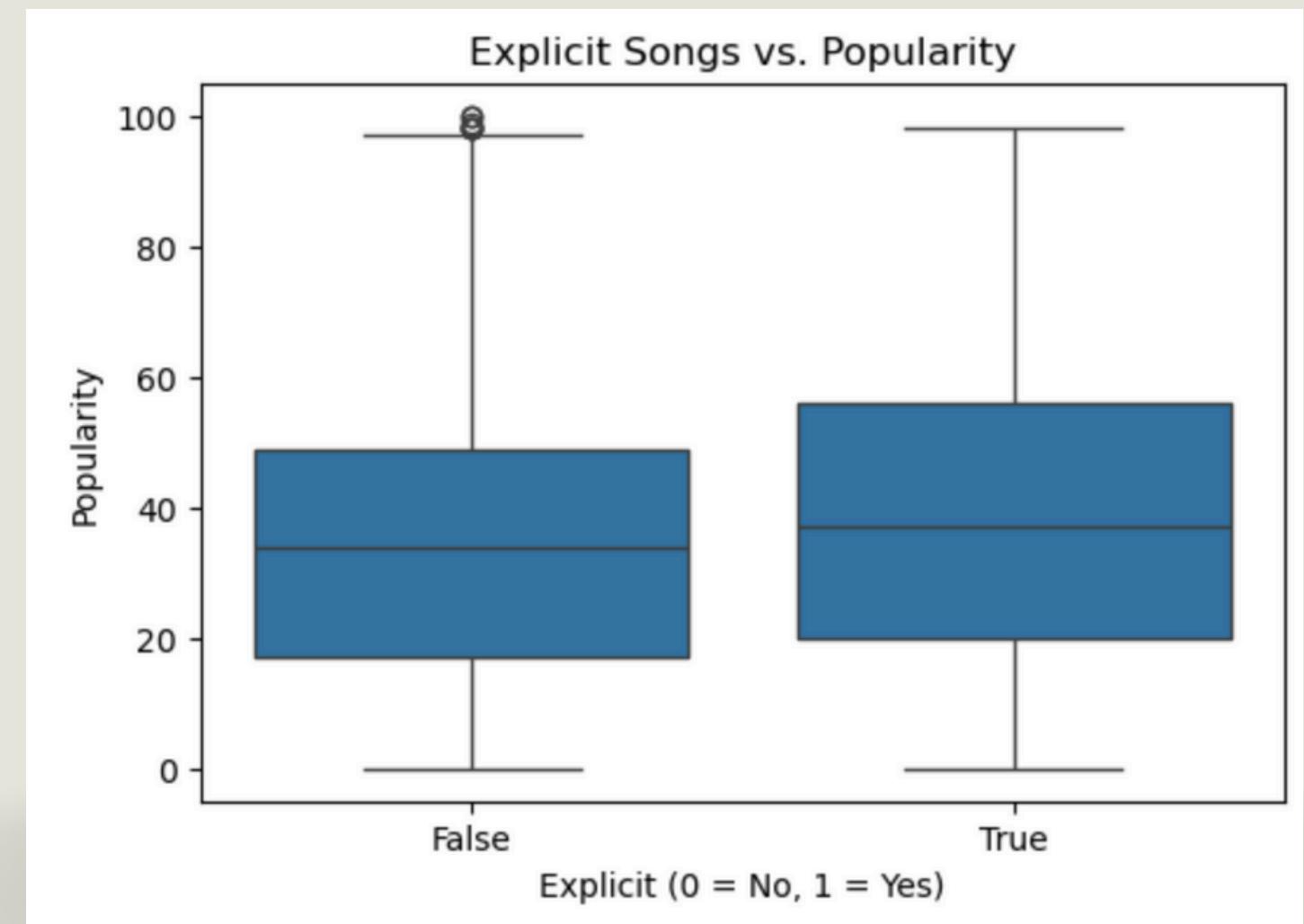
Box plot comparing explicit vs. non-explicit songs

Key Takeaways

- Explicit songs do not have a major advantage in popularity.
- No strong correlation between explicit content and popularity.

Business Insights

- Artists do not need explicit content to create hit songs.
- Explicit lyrics may be more common in certain genres (e.g., hip-hop, R&B) but are not a necessity for mainstream success.



Top 10 Genres by Popularity

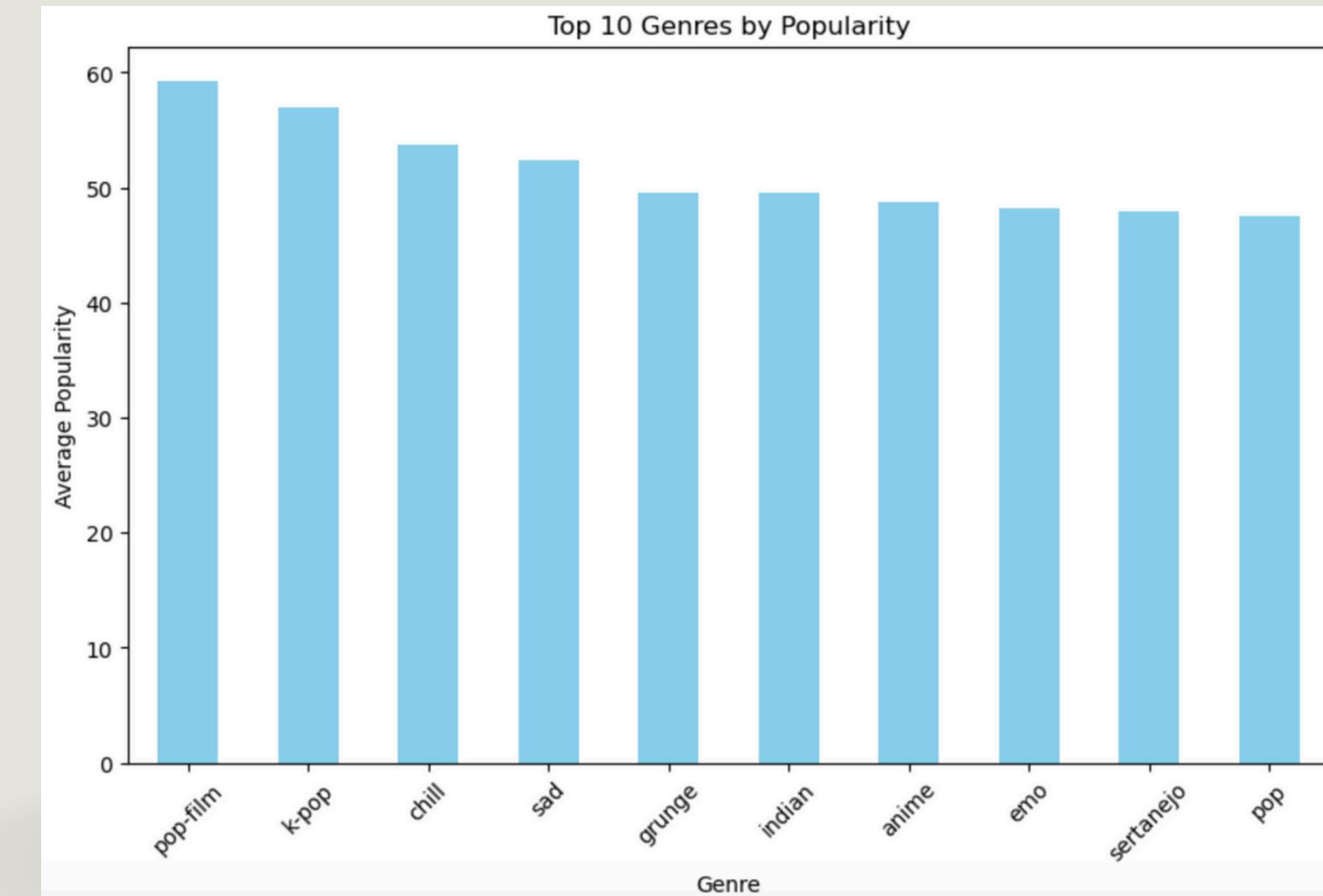
Bar chart of the most popular genres

Key Takeaways

- Pop-film and K-pop are the most popular genres.
- Chill and sad genres are also highly popular.

Business Insights

- Streaming services should focus on playlist curation for these top genres.
- Emerging genres like anime and Indian music are gaining traction, offering growth opportunities.



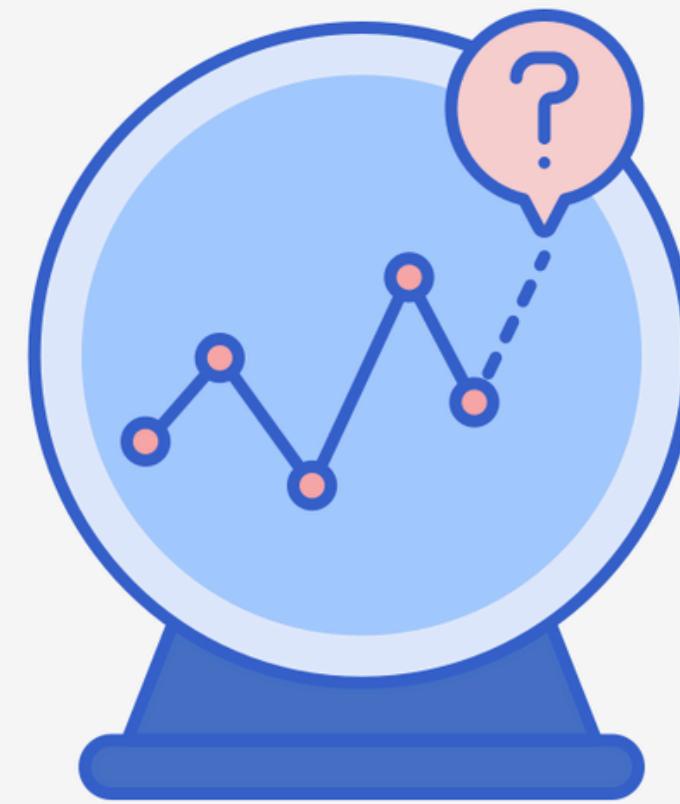
Predictive Model Results

Model Used: Random Forest Regression

Key Result: Mean Absolute Error (MAE)
 ≈ 10.78

Business Insights

- Audio features alone do not fully predict popularity, external factors matter more.
- Streaming services and record labels can leverage predictive models to assess potential hit songs, but marketing strategies remain crucial.
- Future research should explore non-audio factors (e.g., social media engagement, artist collaborations) to refine predictions.



Conclusion



Business Implications

- No single audio feature guarantees success.
- External factors (marketing, artist reputation, playlist placement) influence popularity more than music composition.
- Streaming platforms & labels should focus on genre-based promotion and mood-driven playlists.

Future research ideas

- Incorporate social media trends, artist recognition, and streaming stats for better prediction.

Thank You

