OPTIMIZING MUSIC SUCCESS ON YOUTUBE

An Analytical Approach to Maximizing Views for Artists and Record Labels

republic records

By Saksham Tiwari

Unlocking the Secrets to Maximizing YouTube Views

As the music industry continues to evolve in the digital age, the success of a track is no longer solely determined by the talent of the artist or the production quality. With platforms like YouTube becoming the primary stage for music discovery and consumption, understanding what drives a song to accumulate millions—or even billions—of views has become crucial for artists, producers, and record labels.

This project was born out of a simple yet pressing question: What makes a music video go viral? Is it the danceability of the track? The energy it exudes? Or perhaps the loudness that catches the listener's attention? With millions of tracks available at the listener’s fingertips, what differentiates a hit from a miss?

Imagine a formula that could significantly boost your track's visibility on YouTube.

Content

INTRODUCTION

BUSINESS PROBLEM:

How can artists and sound engineers optimize music quality and reach, while managers and record labels identify tracks to attract the highest views on Youtube?

OBJECTIVE:

TO ANALYZE SONG FEATURES AND PERFORMANCE METRICS TO PREDICT AND ENHANCE THE SUCCESS OF TRACKS, GUIDING STRATEGIC DECISIONS FOR MUSIC VIDEO PRODUCTION.

Data Overview

Source of Dataset:

- The dataset was sourced from Youtube and Spotify containing detailed information on various song features and their performance.

Key Columns:

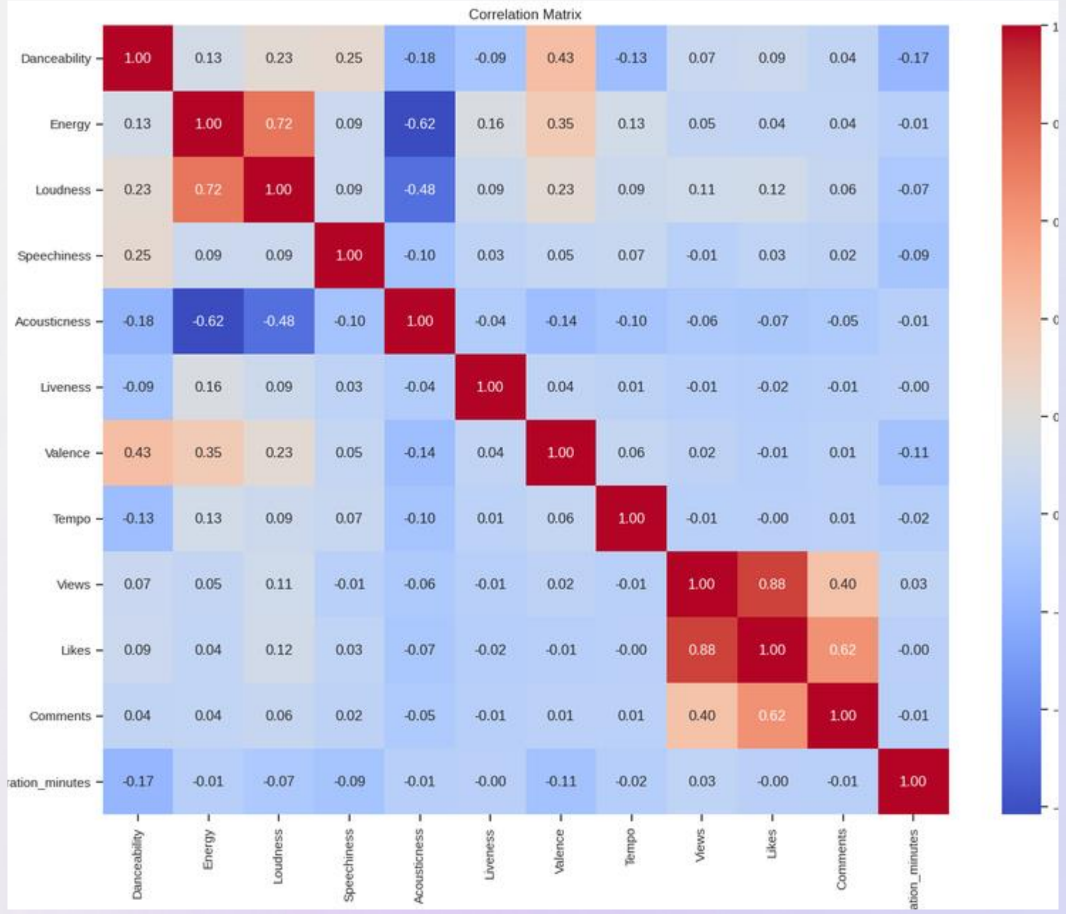
- The dataset originally consisted of 28 columns and 20,718 rows.
- Key features included song attributes like Danceability, Energy, Loudness, Speechiness, Acousticness, Liveness, Valence, and Tempo.
- Performance metrics such as Views, Likes, Comments, and Duration_minutes were also included.

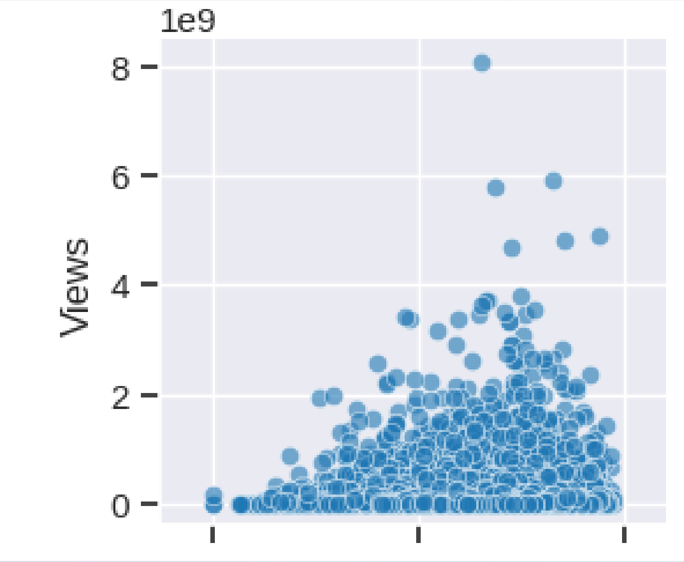
Exclusions Made:

- Columns excluded due to redundancy, irrelevance, or high correlation include Stream, Likes, Channel, Artist, Album, Track, Comments, and Title.
- Instrumentalness was removed due to its skewed distribution.
- Tracks with less than 1 million views were excluded to focus on songs with significant impact.

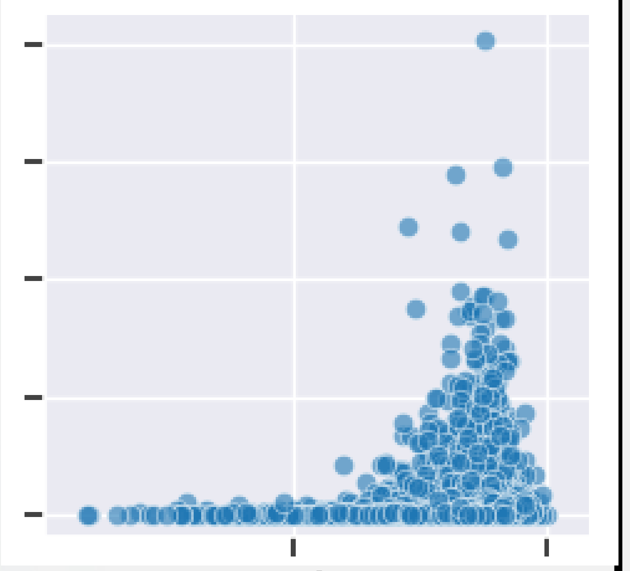
Initial Insights

Popular tracks (high views) often have higher engagement metrics (Likes, Comments) and are characterized by higher Energy, Loudness, and Danceability.



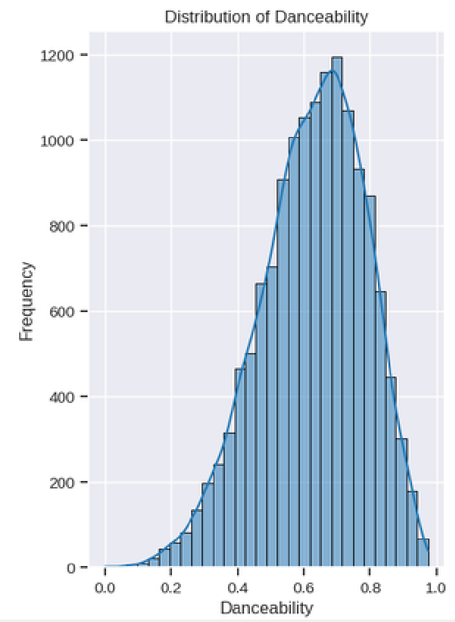


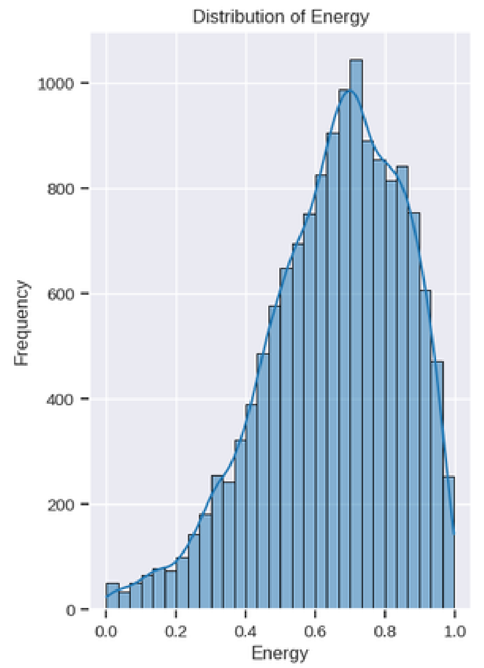
Danceability

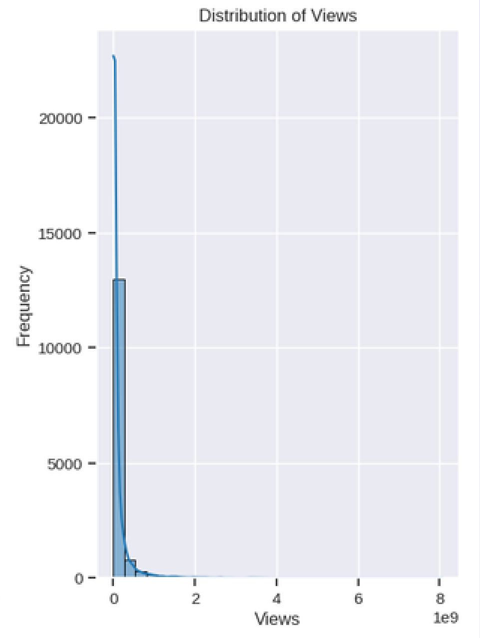


Loudnes

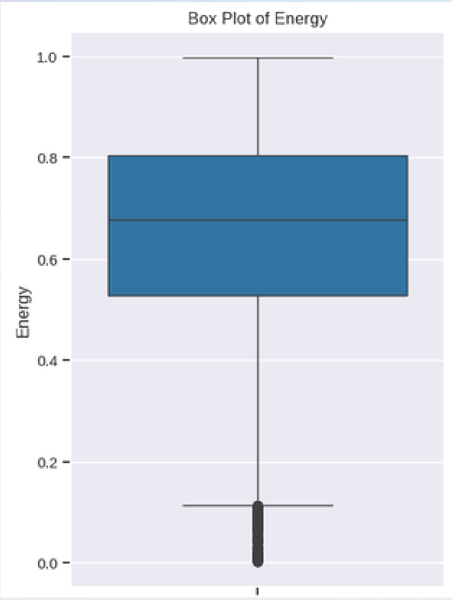
Clusters show varying popularity across different levels of Loudness and Danceability, hinting at sweet spots for track attributes.

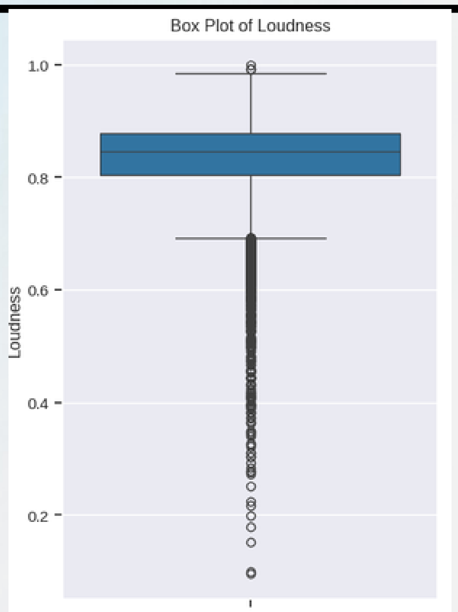


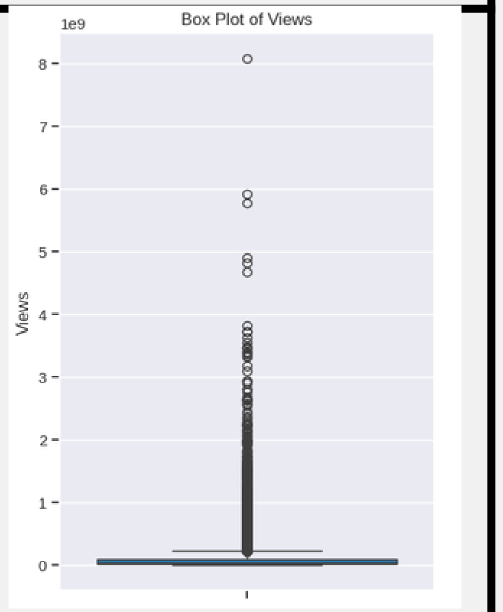




Views are heavily skewed, with most tracks clustering between 1 and 100 million views. Energy and Danceability have more balanced distributions, indicating their importance across the board



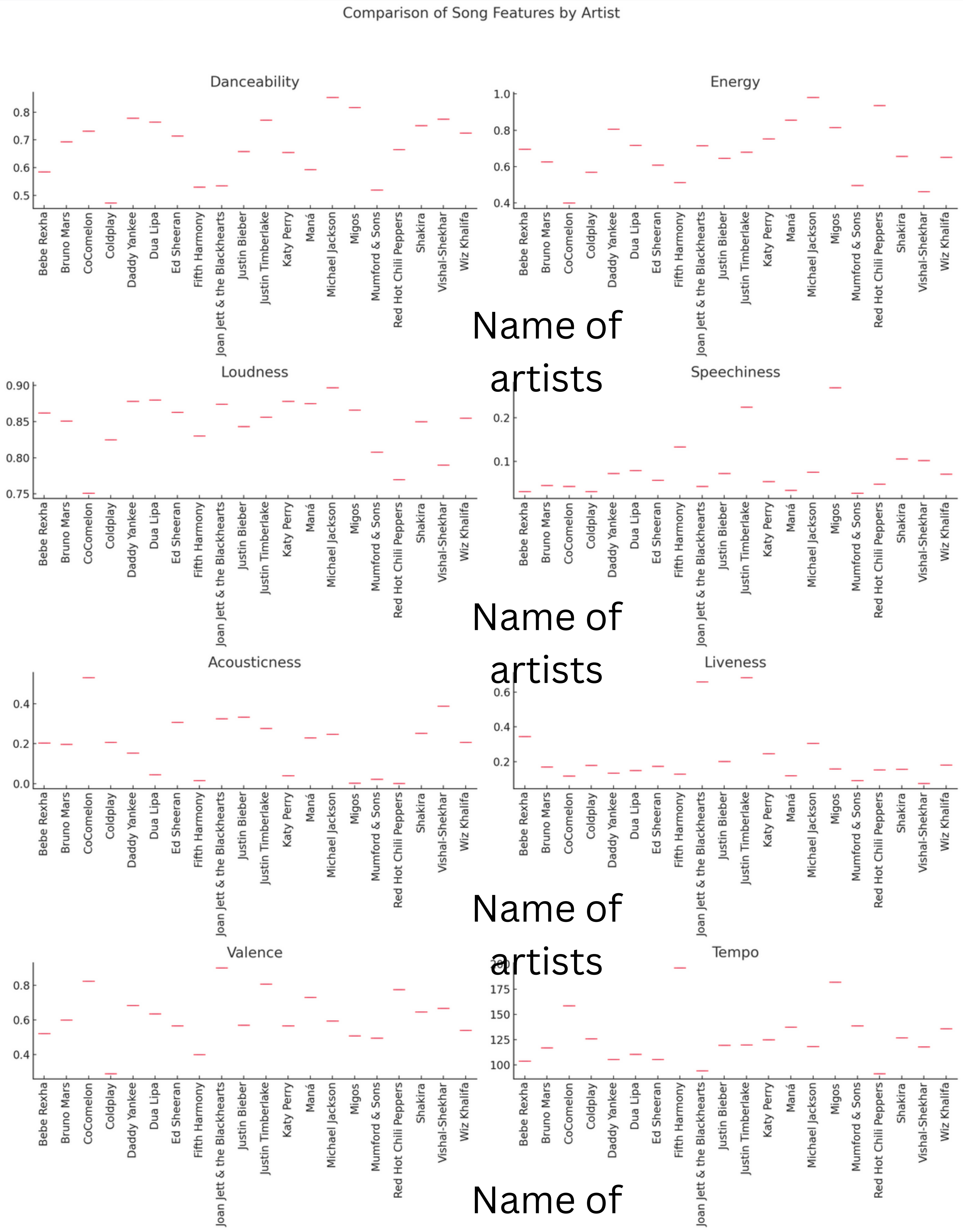




Consistent levels of Loudness and Energy in popular tracks, with certain songs not following the normal trend.

We compared the **top 10 artists** from the 1M to maximum views group with those from the 1M to 100M views group to understand the key characteristics that differentiate the top-tier artists.

SO WHAT DID WE DISCOVER???



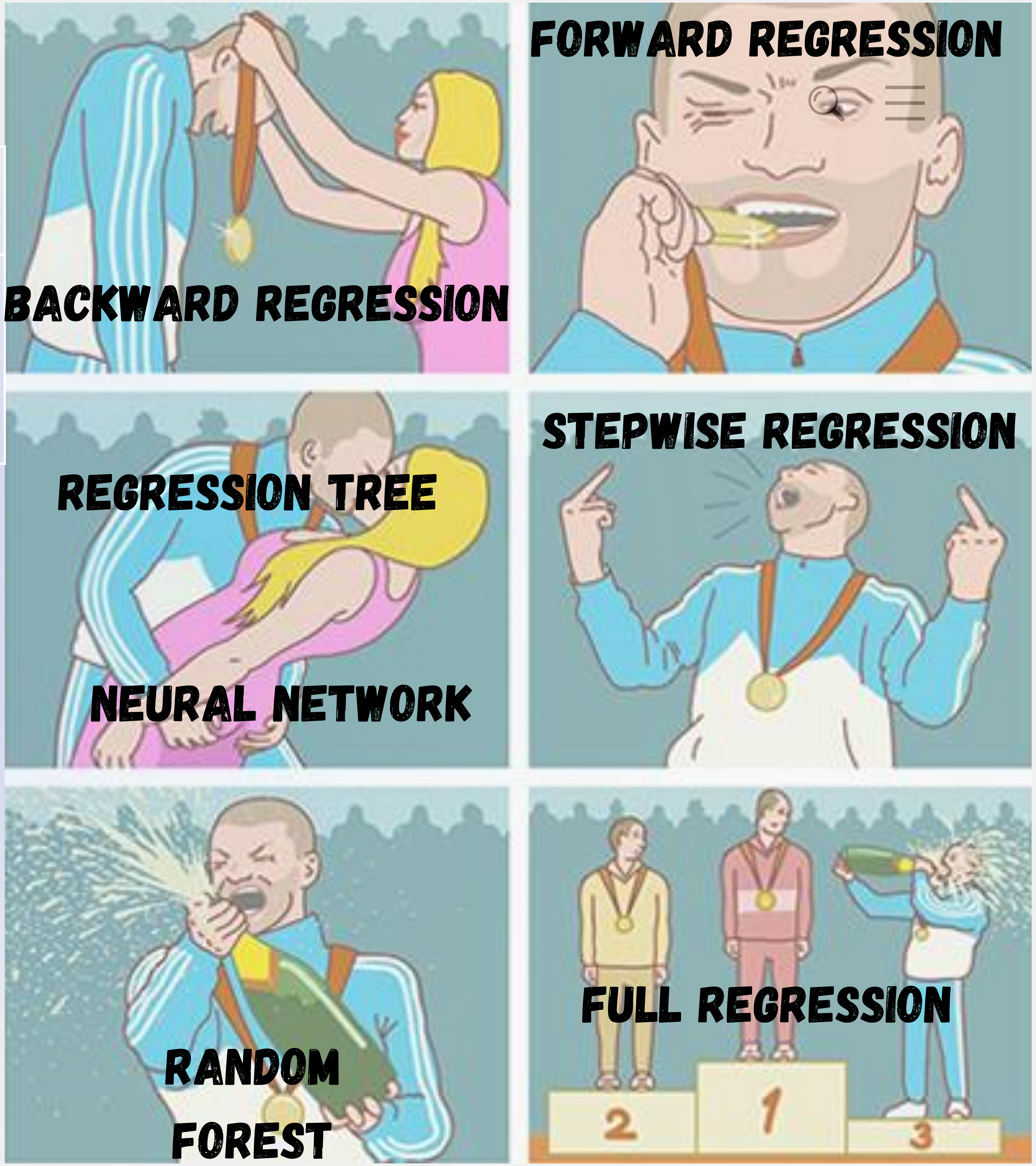
WE SAW THAT SONG VIDEO ACHIEVING HIGH VIEWS TEND TO HAVE

- HIGH DANCEABILITY TREND
- LOUDNESS IS ON THE UPPER BAND AS WELL.
- LIVENESS IS LOW.
- THE TEMPO HOVERS AROUND 125 TO 100 BAND.

Modelling Result

Model	Mean Error (ME)	RMSE	MAE	MPE (%)	MAPE (%)
Backward Regression	581,077.05 78	24,625,953. 6892	19,528,362.4278	-273.9111	306.9864

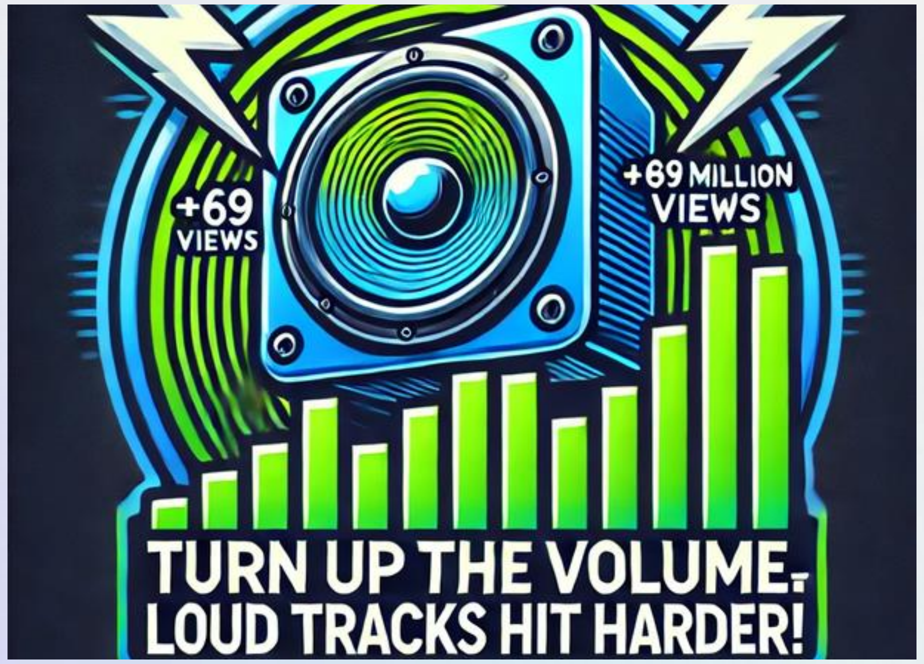
- The Backward Regression model was chosen because it balances accuracy and simplicity.
- It performs comparably to other models but with fewer variables and reducing complexity.
- This model highlights the most important factors that influence YouTube views, making it a practical tool for artists and record labels to optimize their strategies.



SO WHAT DID THE BACKWARD REGRESSION ACTUALLY SAY??

1. Danceability: The Beat That Moves the Crowd

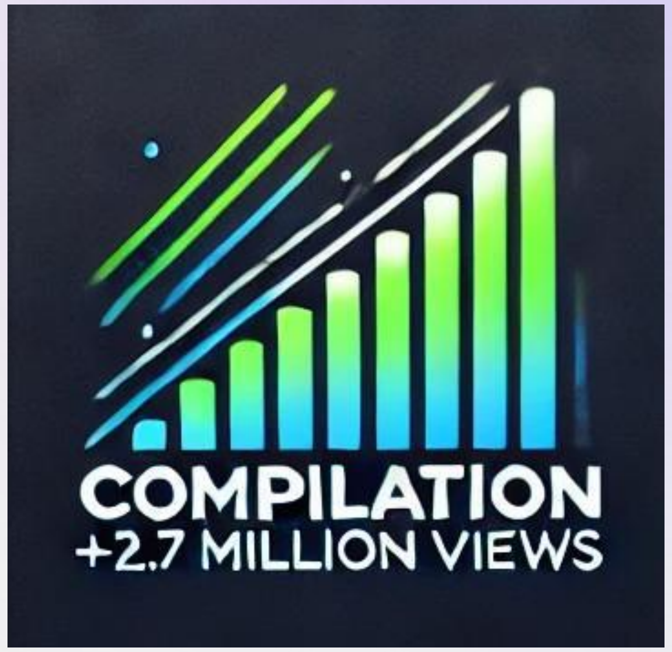
2. Energy: Finding the Sweet Spot



3. Loudness: Make Some Noise

4. Tempo: The Pace That Sets the Tone

5. Track Length: Give Them a Little More



6. Album Strategy: Compilation vs. Singles

7. Licensing: The Power of Protection +4.9 million view

8. Official Videos: The Visual Advantage

9. Key to the song

Recommendations

- **Licensing Matters:** Ensure all tracks are licensed as this can significantly boost visibility and views.
- **Create Official Videos:** Investing in official music videos is crucial as they greatly enhance the track’s potential reach and popularity.
- **Focus on Danceability and Loudness:** Tracks that are more danceable and have balanced loudness levels tend to perform better in terms of views.
- **Strategic Album Placement**: Releasing tracks as part of compilation albums rather than singles can help achieve better visibility and more views.
- **Optimize Track Duration:** While typical track lengths are around 3 to 5 minutes, slightly longer tracks can engage audiences better, leading to higher views.
- **Leverage Popular Keys:** Consider using musical keys that are more commonly associated with higher views, such as C major or A minor.

NEXT STEPS

- To expand on the scope of this project. Let’s bring in more variables and factors affecting the views on a song.
- Different blocks of views can be created for a more focused analysis for example from 10k views to 100k view, 100k to 500k views and so on. These smaller blocks will give a deeper insights.
- The trends in songs are always evolving so keeping the dataset of songs has to be kept in mind.
- Not sticking to just one model for all the insight. As dynamic the songs are let’s involve different models to uncover the patterns associated with a song.
- Build a strong analytics team for artists who want to work in a specific genre.

Thank You

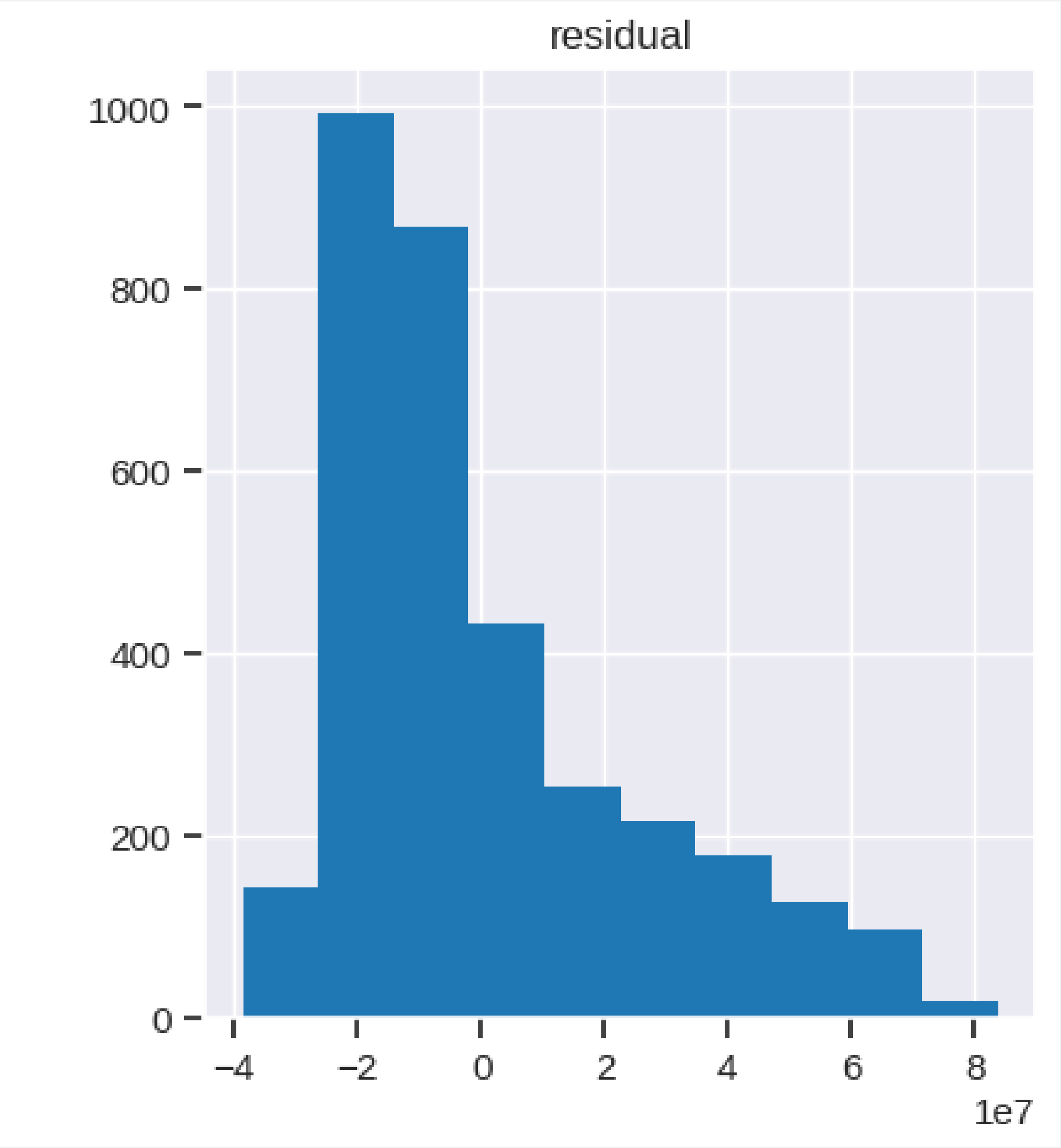
APPENDIX

1. Model Comparison Table

Model	Mean Error (ME)	RMSE	MAE	MPE (%)	MAPE (%)
Full Regression	566,334.7585	24,623,324.0997	19,526,341.1532	-274.3988	307.5128
Backward Regression	581,077.0578	24,625,953.6892	19,528,362.4278	-273.9111	306.9864
Forward Regression	581,077.0578	24,625,953.6892	19,528,362.4278	-273.9111	306.9864
Stepwise Regression	581,077.0578	24,625,953.6892	19,528,362.4278	-273.9111	306.9864
Decision Tree	552,495.9803	24,652,330.9347	19,539,825.2468	-273.7849	306.5290
Random Forest	-750,759.0322	24,854,038.0771	19,935,836.9468	-291.4426	323.1347
Neural Network	-	24,859,165.8217	19,825,625.4959	-	-

2. Odd Ratios from Backward Elimination

Predictor	Coefficient
Danceability	+9,010,455
Energy	-8,625,582
Loudness	+69,262,870
Tempo	+16,609
Duration (minutes)	+2,629,623
Flag Outlier in Loudness	+3,984,852
Album Type (Compilation)	+2,694,618
Album Type (Single)	-3,367,937
Key (C Major/A Minor)	+2,247,446
Licensed (True)	+4,874,986
Official Video (True)	+6,075,166



On valid

3. Regressor tree