AJEENKYA
D Y PATIL UNIVERSITY

Newton School of Technology

ZIPF'S LAW PROJECT

Team - Insighters

ANUJ || AMRIT || NAKUL || SAUBHAGYA

Subject - Fundamental Of AI

TEAM MEMBERS

ANUJ CHHAJED
anuj.chhajed@adypu.edu.in

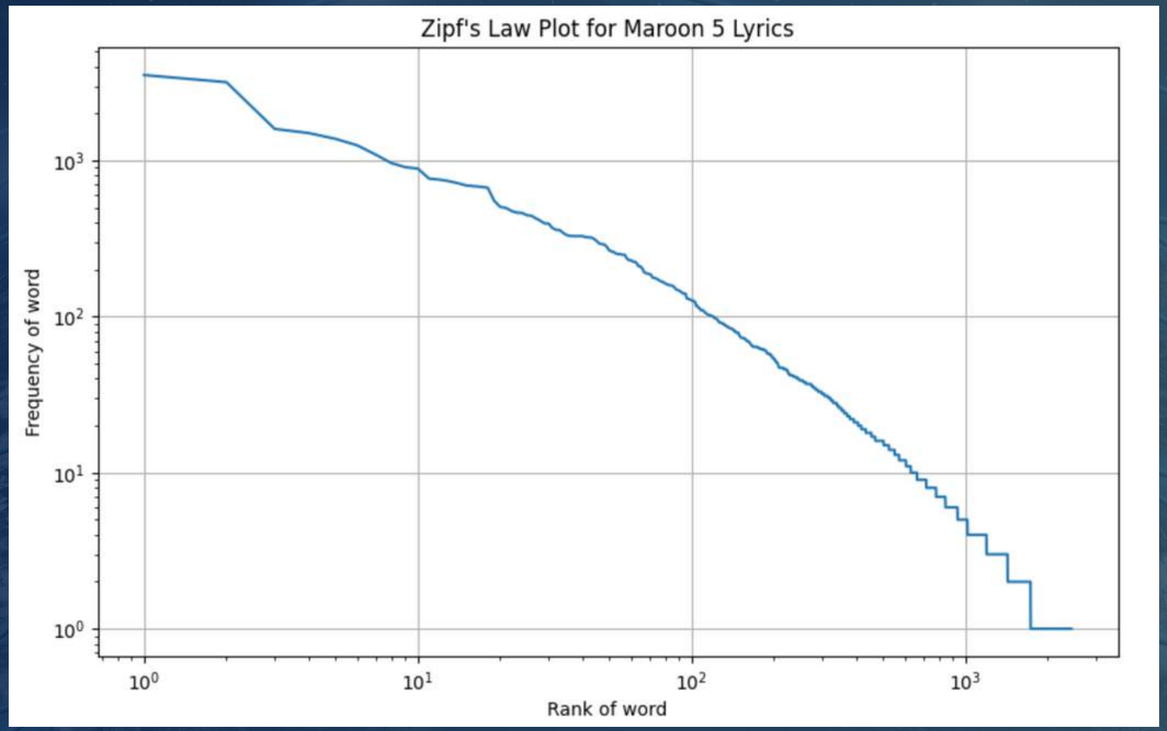
AMRIT KUMAR MAHTO
amrit.mahto@adypu.edu.in

NAKUL SHARMA
nakul.sharma@adypu.edu.in

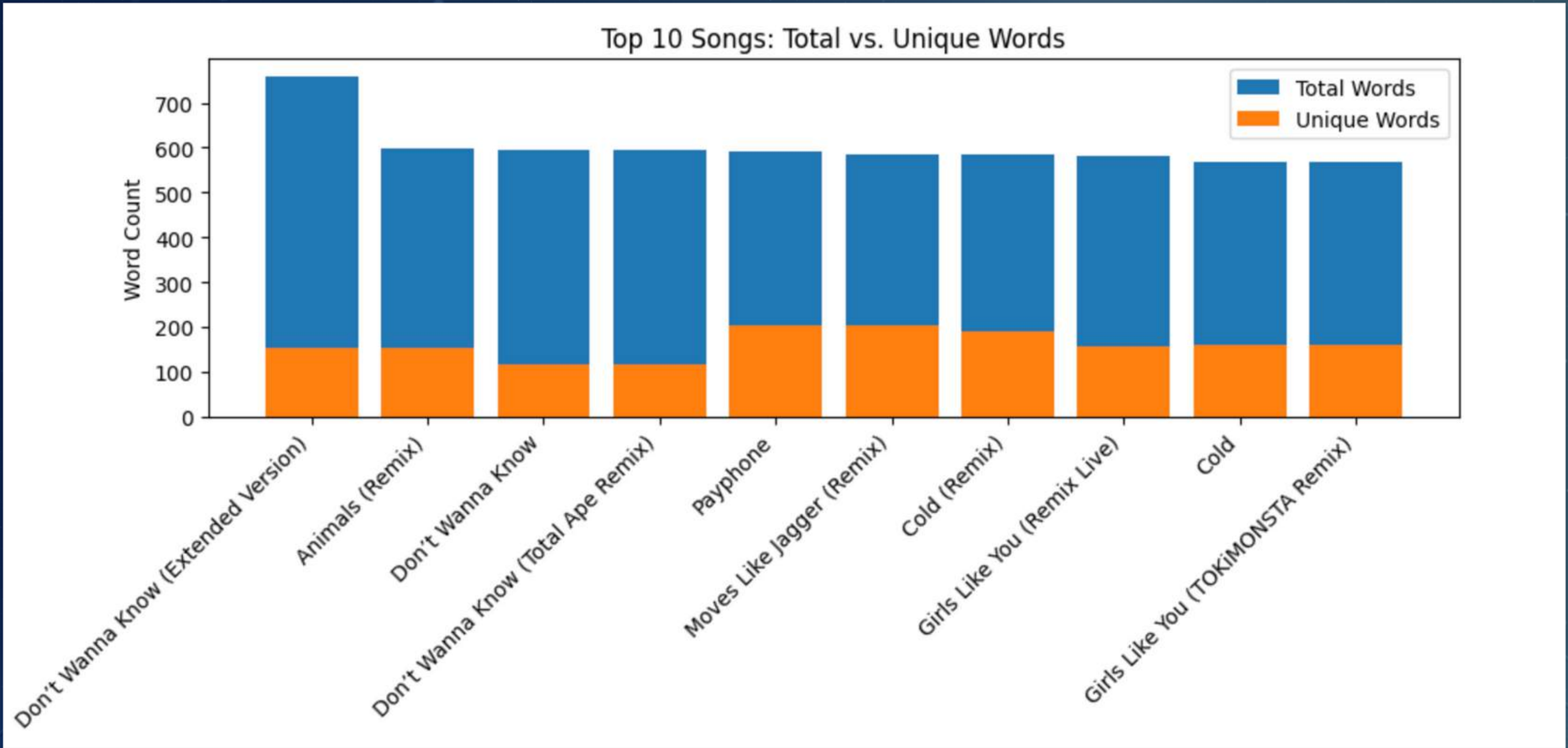
SAUBHAGYA ANUBHAV
Saubhagya.Anubhav@adypu.edu.in

UNDERSTANDING ZIPE'S LAW

We analyzed the word frequencies in Maroon 5's lyrics and plotted them against their ranks on a log-log scale. The resulting graph showed a nearly straight downward line, which matches the expected pattern of Zipf's Law. This indicates that the most common words appear far more frequently than rarer ones, following the inverse rank-frequency relationship. Hence, the lyrics dataset supports the presence of Zipf's Law.

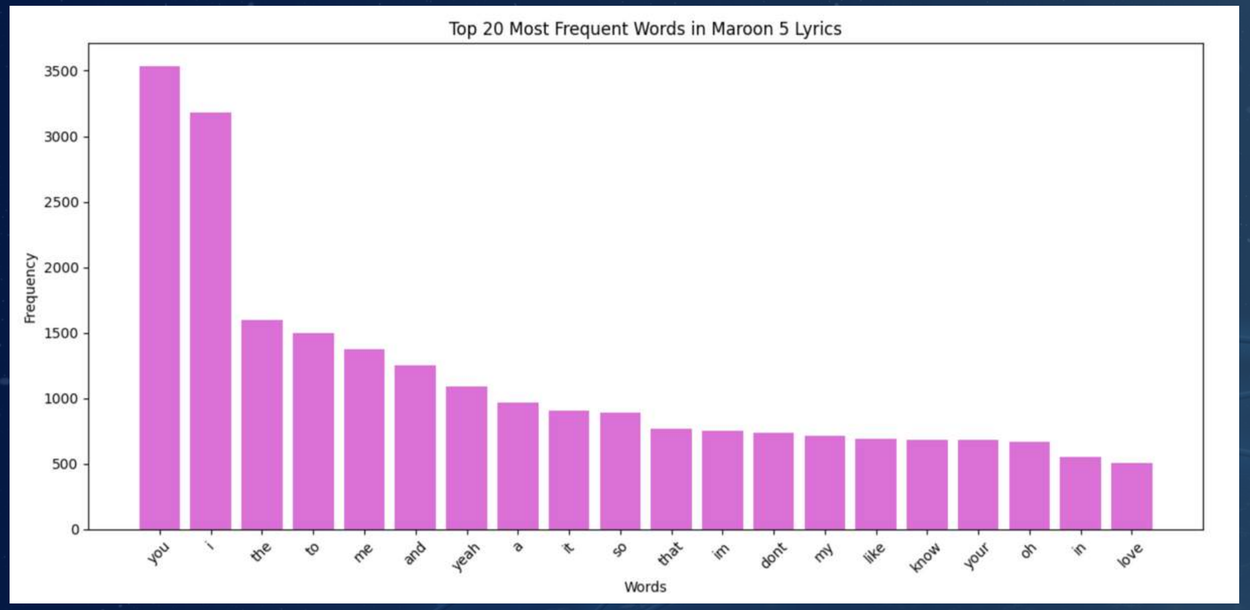


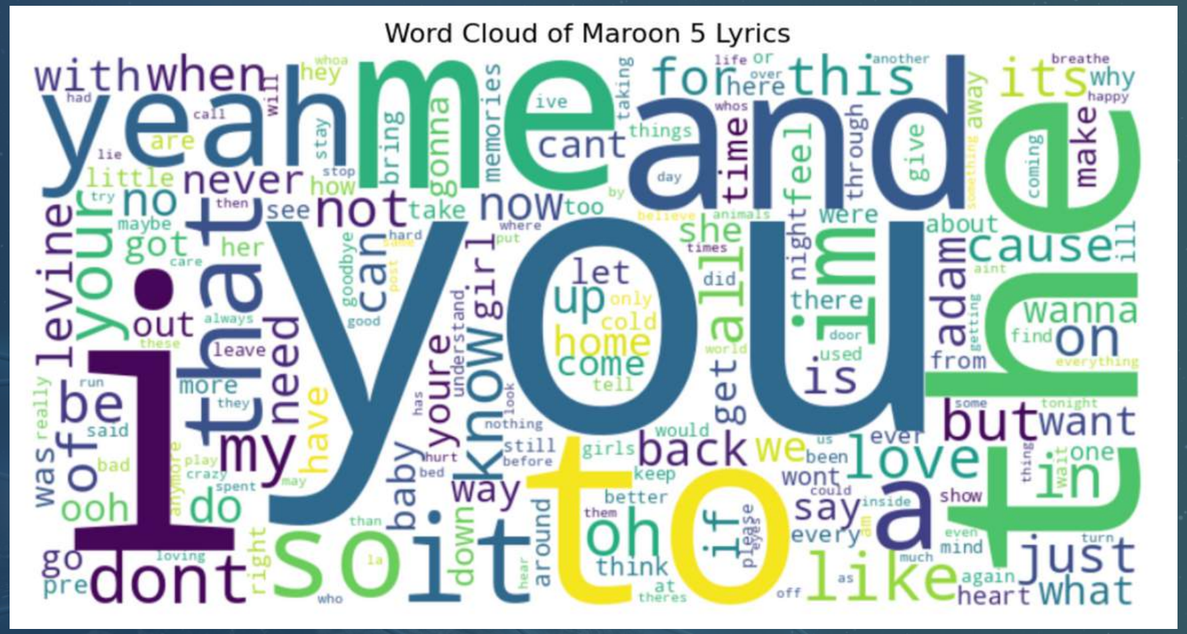
ANUJ'S CONTRIBUTION



THIS BAR CHART SHOWS THE TOP 10 MAROON 5 SONGS BASED ON TOTAL WORD COUNT. IT COMPARES THE TOTAL NUMBER OF WORDS WITH THE NUMBER OF UNIQUE WORDS IN EACH SONG, HELPING US UNDERSTAND THE LYRICAL LENGTH AND VOCABULARY RICHNESS USED IN DIFFERENT SONGS.

ANUJ'S CONTRIBUTION

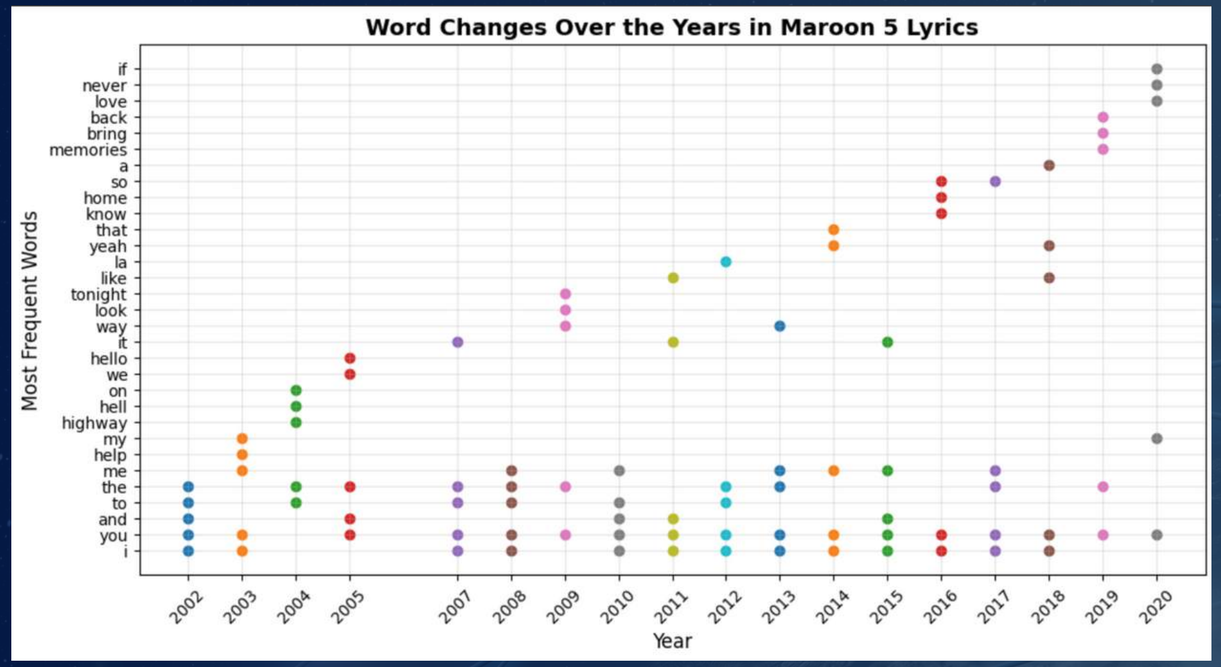




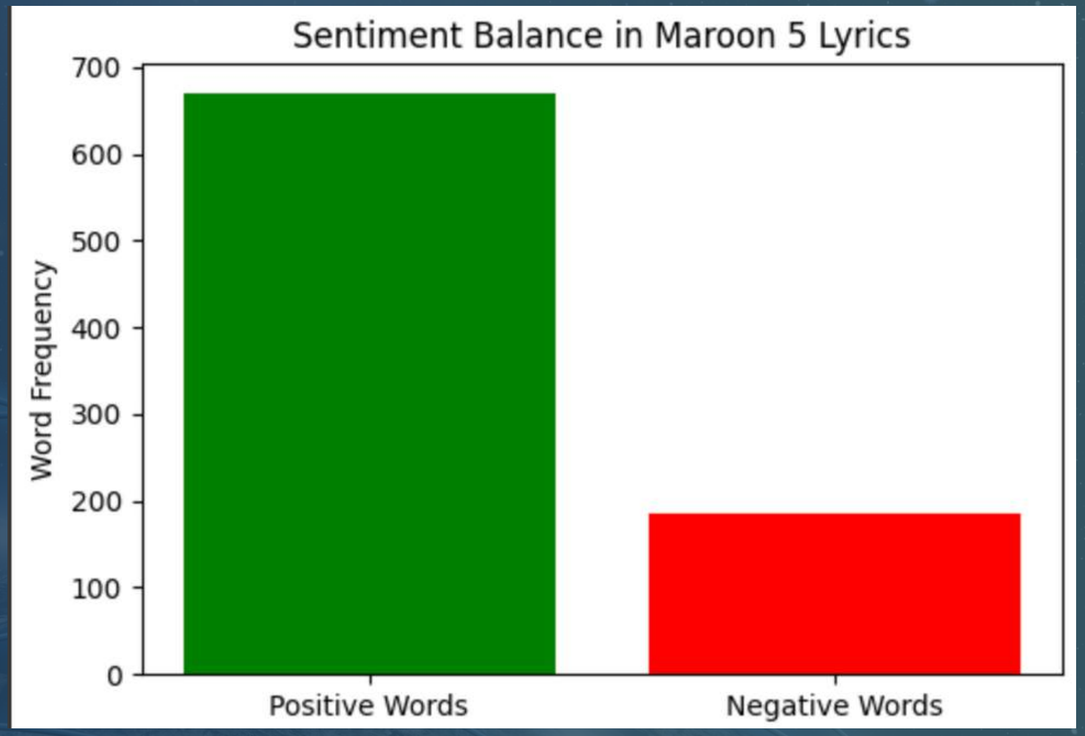
THIS BAR CHART DISPLAYS THE 20 MOST FREQUENTLY USED WORDS IN MAROON 5'S LYRICS, HIGHLIGHTING COMMON THEMES AND REPEATED VOCABULARY.

THIS WORD CLOUD VISUALLY REPRESENTS THE MOST FREQUENT WORDS IN MAROON 5'S LYRICS — LARGER WORDS APPEAR MORE OFTEN, OFFERING AN INTUITIVE VIEW OF LYRICAL THEMES.

AMRIT'S CONTRIBUTION

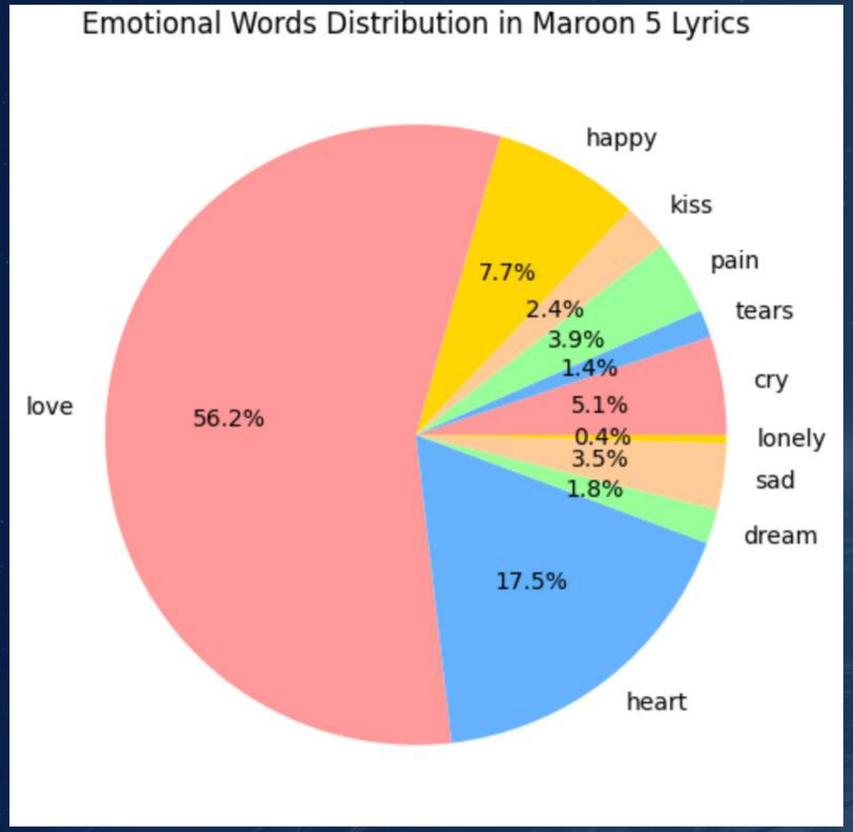


TRACKING THE MOST FREQUENTLY USED WORDS IN LYRICS ACROSS DIFFERENT YEARS AND VISUALIZING THEIR CHANGES OVER TIME.

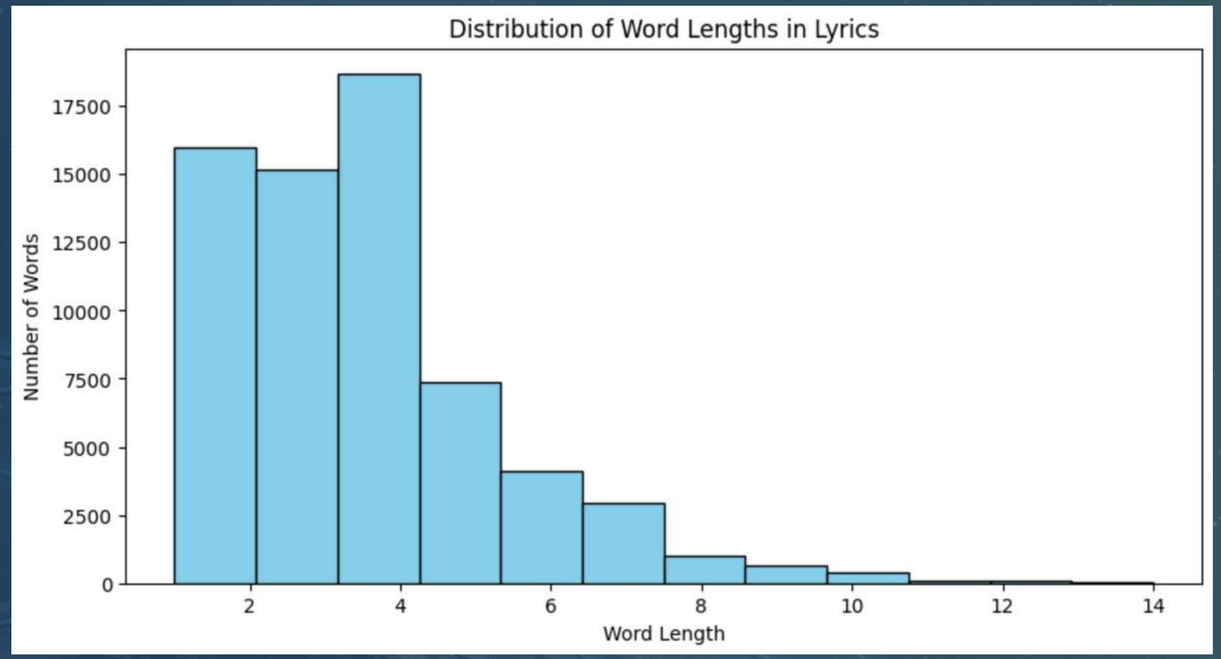


ANALYZING THE SENTIMENT BALANCE IN LYRICS BY COUNTING THE OCCURRENCES OF POSITIVE AND NEGATIVE WORDS

NAKUL'S CONTRIBUTION

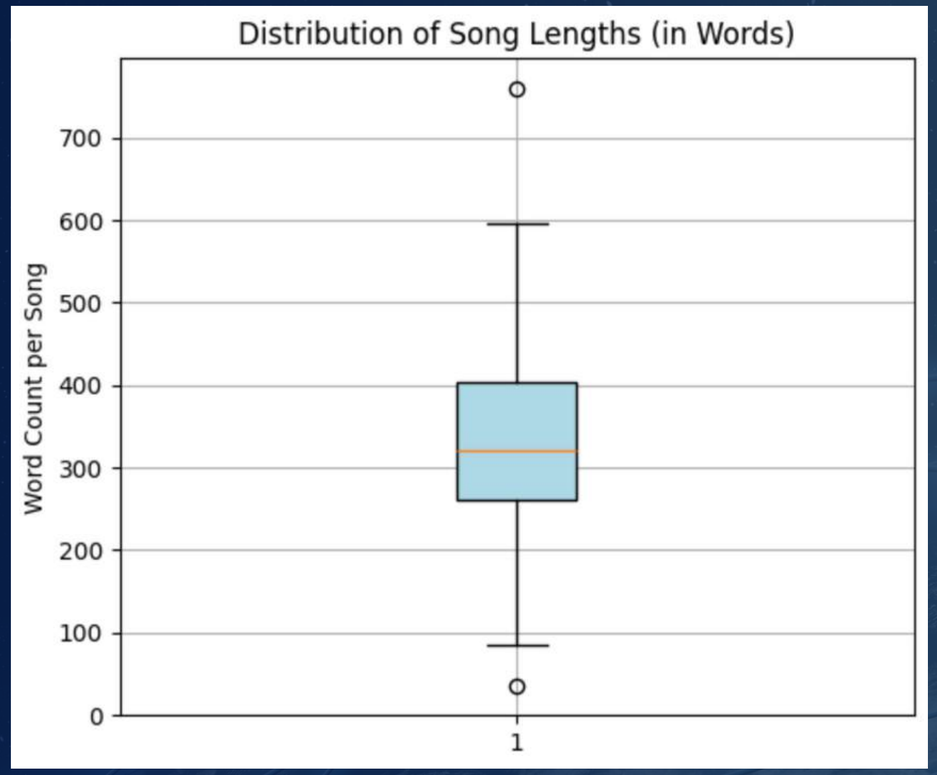


VISUALIZING THE DISTRIBUTION OF EMOTIONAL WORDS IN LYRICS USING A PIE CHART TO UNDERSTAND THE PREVALENCE OF DIFFERENT EMOTIONS IN THEIR SONGS.

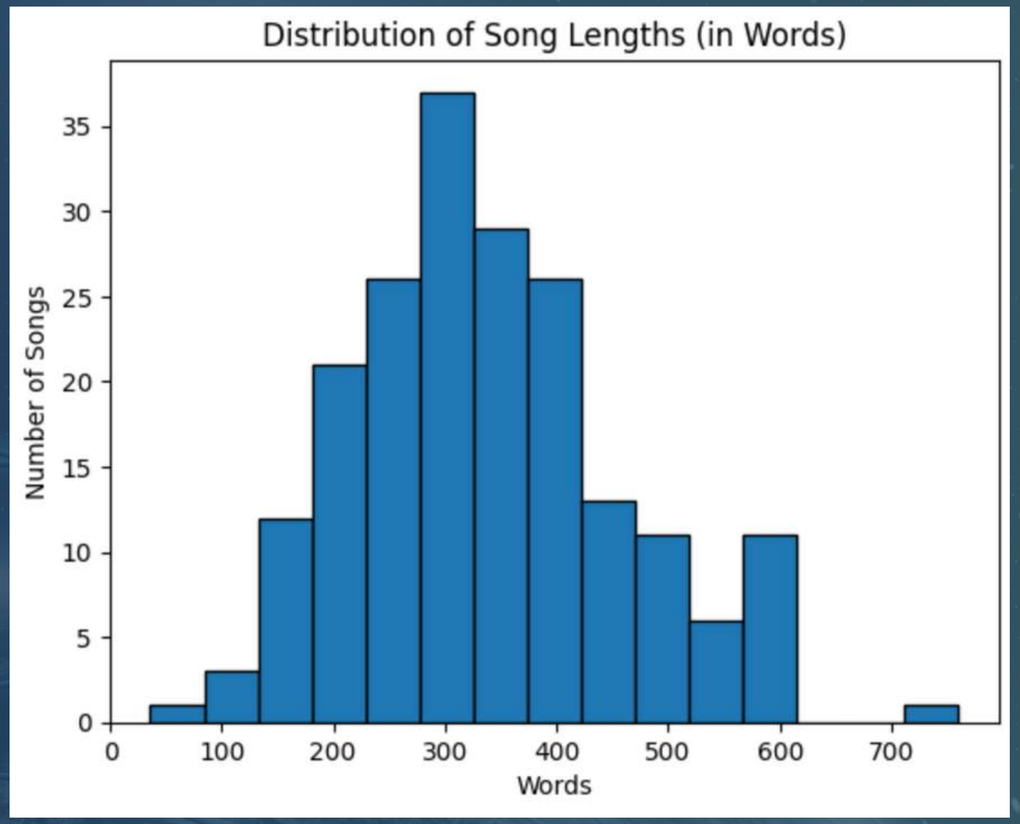


ANALYZING THE DISTRIBUTION OF WORD LENGTHS IN THE LYRICS, VISUALIZING HOW FREQUENTLY DIFFERENT WORD LENGTHS APPEAR USING A HISTOGRAM.

SAUBHAGYA'S CONTRIBUTION



DISTRIBUTION OF SONG LENGTHS IN TERMS OF WORD COUNT, VISUALIZING HOW WORD COUNT VARIES ACROSS DIFFERENT SONGS USING A BOX PLOT.



ANALYZED THE DISTRIBUTION OF SONG LENGTHS BASED ON WORD COUNT, VISUALIZING HOW SONG LENGTHS VARY USING A HISTOGRAM.

FINDINGS

1. Zipf's Law Observed

The log-log plot of word frequency vs. rank formed a near-linear downward curve, confirming that Maroon 5 lyrics follow Zipf's Law — a few words appear very frequently, while many appear rarely.

2. Dominance of Common Words

The top 20 most frequent words include simple, emotionally charged terms like love, baby, and heart, showing lyrical themes are often romantic and repetitive.

3. Variation Across Songs

Some songs have significantly higher total word counts than others, indicating lyrical density differs across tracks. A few songs also have a high number of unique words, suggesting lyrical richness.

4. Positive vs. Negative Emotion

Sentiment-related words showed a greater number of positive terms like love, smile, and beautiful than negative ones, reflecting the overall upbeat tone of the band's music.

5.Top 10 Songs by Word Usage

Analyzed total and unique word counts for each song. Some songs like 'Sugar' and 'Animals' had a high total word count but relatively lower unique word count — suggesting repetitive phrasing.

THANK YOU!

Team - Insighters