Post Malone Songs Analysis

Team Name: Sai Teja Ki Team

Members:

Aayush Chaturvedi - Aayush.Chaturvedi@adypu.edu.in

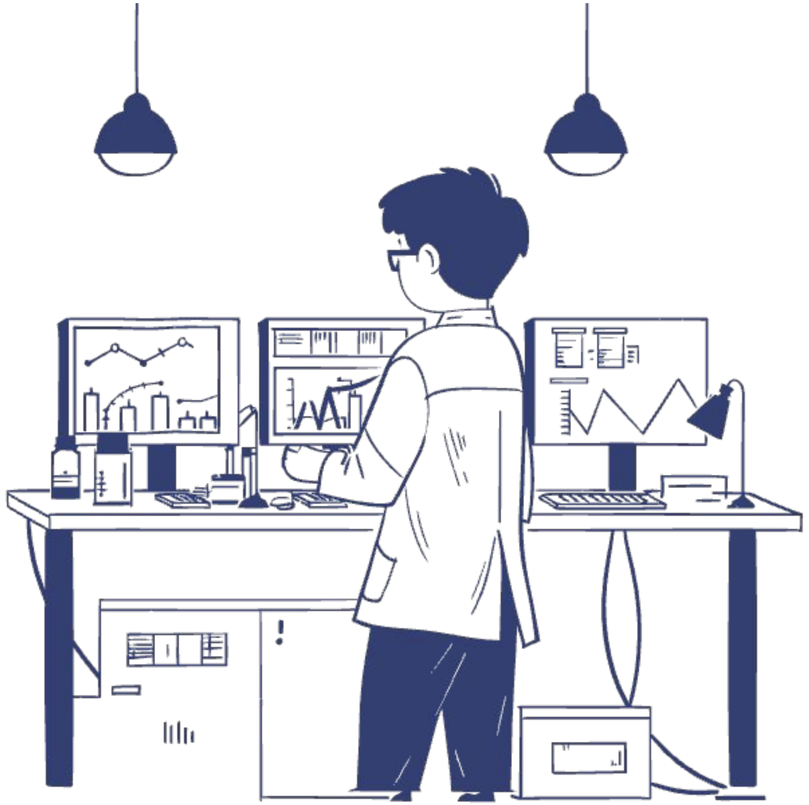
Veeramalla Saiteja - veeramalla.saiteja@adypu.edu.in

Ashvin Tiwari - Ashvin.K@adypu.edu.in

Abhinav Vinod Bhusagare - Abhinav.Bhusagare@adypu.edu.in

Analysis Of PostMalone Data Set

This presentation explores a detailed analysis of Post Malone songs dataset, covering metadata such as song titles, albums, release years, and song lyrics. The analysis includes various statistical and linguistic examinations to uncover patterns in release timing, lexical diversity, and word usage. The dataset contains 148 entries with fields on song title, album, year, release date, and lyrics.



Dataset Overview and Zipf's Law Concept

Descriptive Statistics

The dataset features 148 songs spanning 7 albums over 8 years, from 2013 to 2020. Basic summaries include year distribution averages and unique counts of songs and albums, providing foundational insight into data scope and balance.

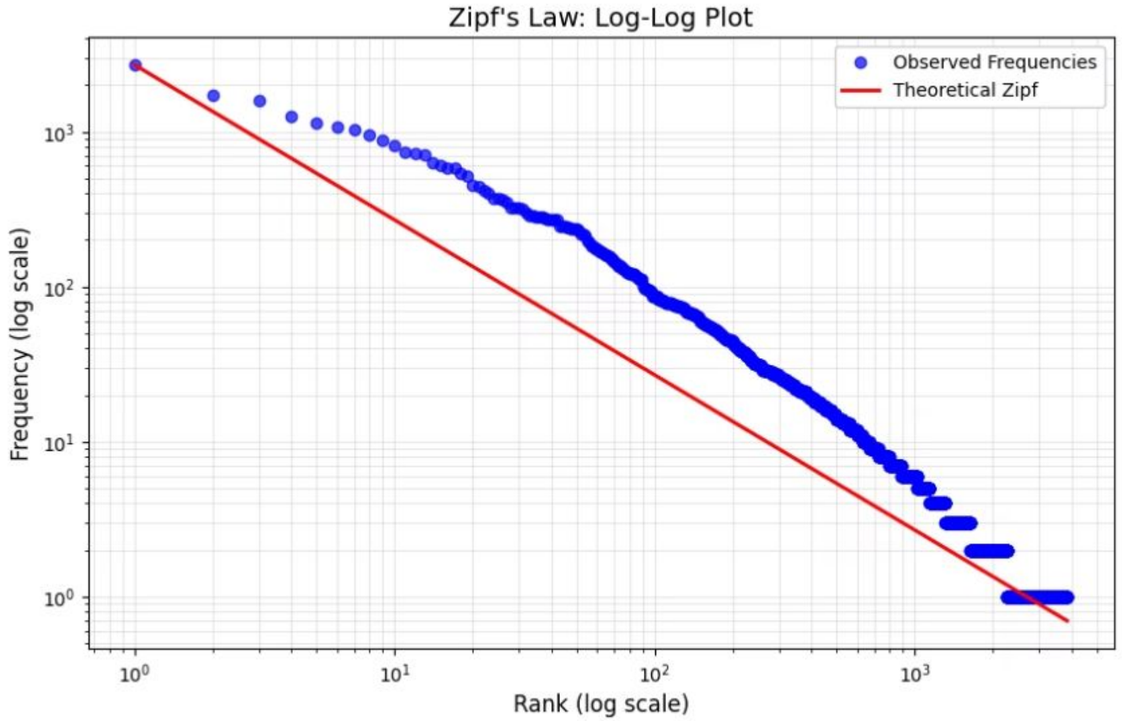
Zipf's Law Analysis

Zipf's Law states that in natural language the frequency of any word is inversely proportional to its rank in the frequency table. The analysis cleans lyrics text, counts word occurrences, and ranks them to assess conformance with this linguistic law.

Zipf’s Law Visualizations

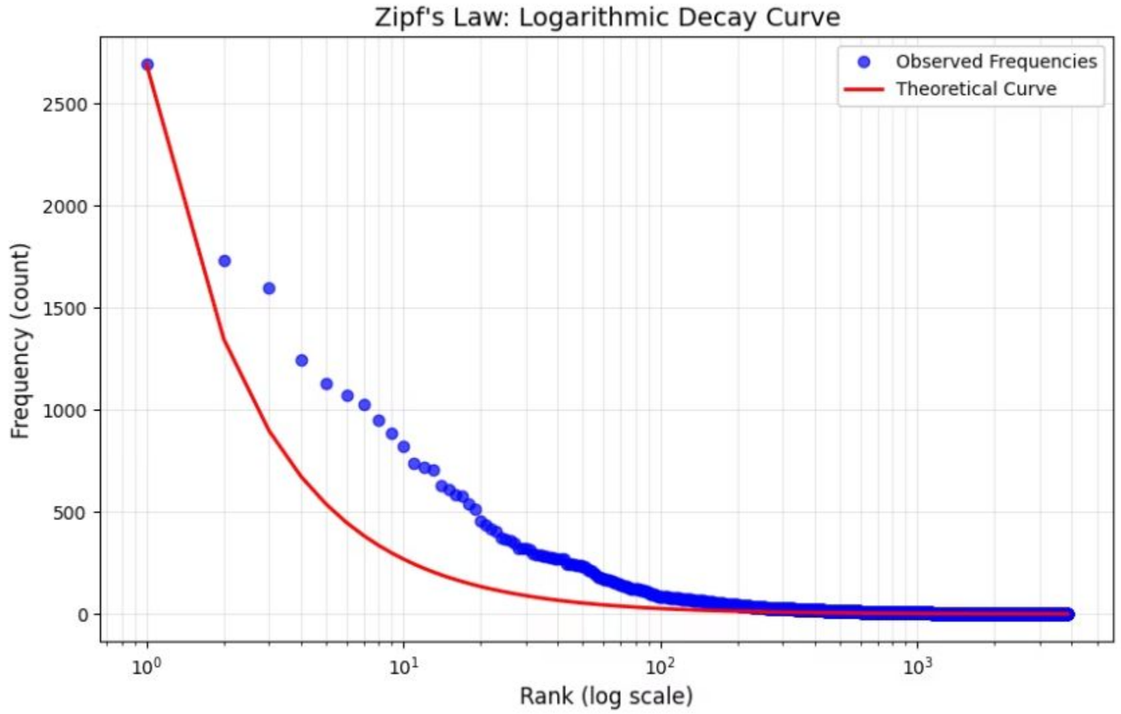
Log-Log Plot

The log-log plot compares observed word frequencies with the theoretical Zipf distribution. Points of observed word rank versus frequency align closely with the expected inverse curve, validating typical language usage patterns within lyrics.



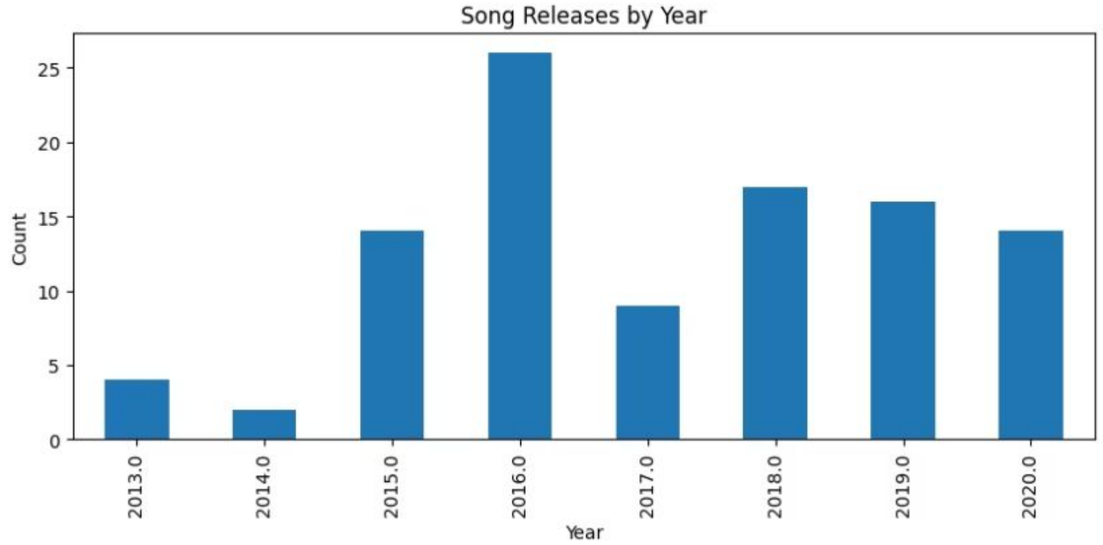
Logarithmic Decay Curve

A linear-scale frequency plot with logarithmic x-axis rank highlights how high-ranking words dominate counts, while lower ranks sharply diminish, illustrating linguistic economy in lyric composition.



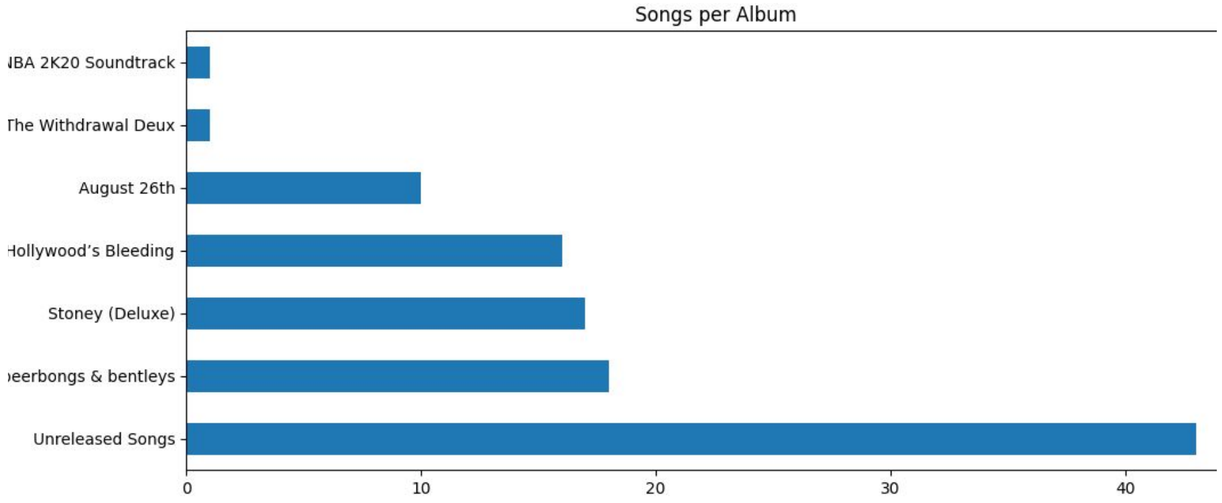
Distribution of Song Releases over Time and Albums

Yearly Release Trends



The bar chart demonstrates fluctuations and peaks in song releases across years. This temporal spread reveals major activity phases in Post Malone's career, especially around albums released from 2015 to 2019.

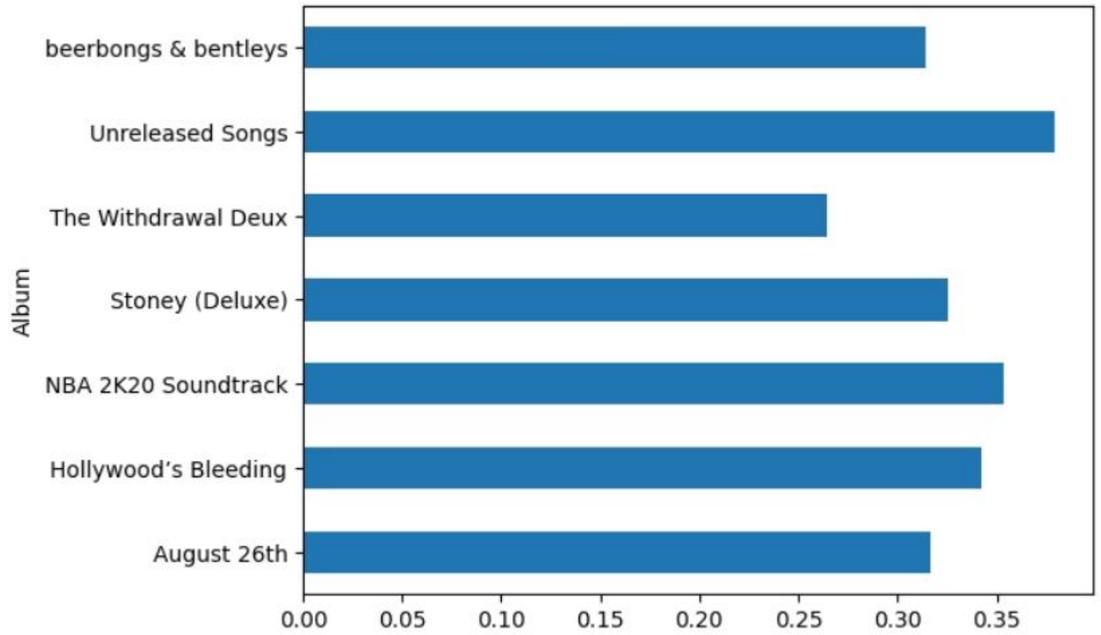
Album-wise Song Counts



Songs per album vary significantly, indicating artistic focus and output intensity. This distribution aids understanding of production emphasis and album impact in dataset composition.

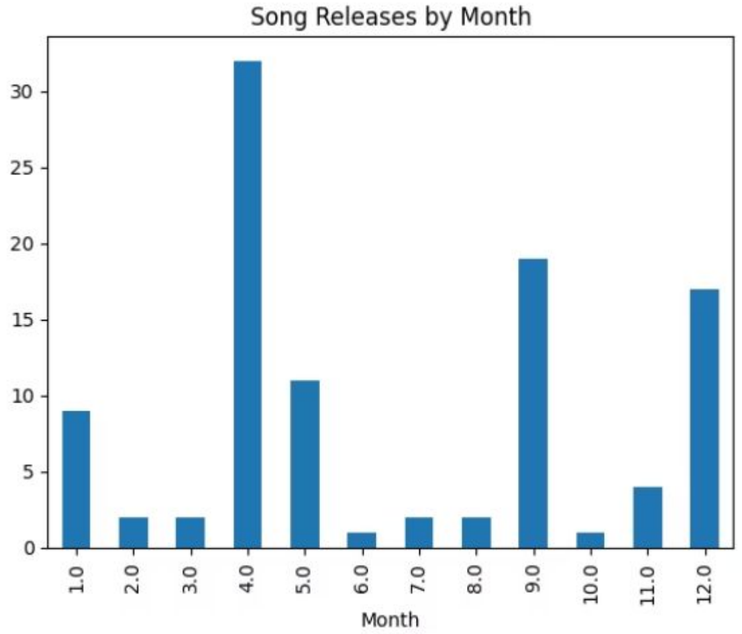
Lexical Diversity and Preferred Release Months

Vocabulary Richness



Analyzing lexical diversity per album quantifies vocabulary variety as a measure of lyric complexity. Variations among albums suggest evolving songwriting style and thematic diversity.

Song Releases by Month



Song releases are not uniformly distributed monthly, indicating possible seasonal patterns in production or marketing strategies. Months with peak releases may correlate with promotional cycles.

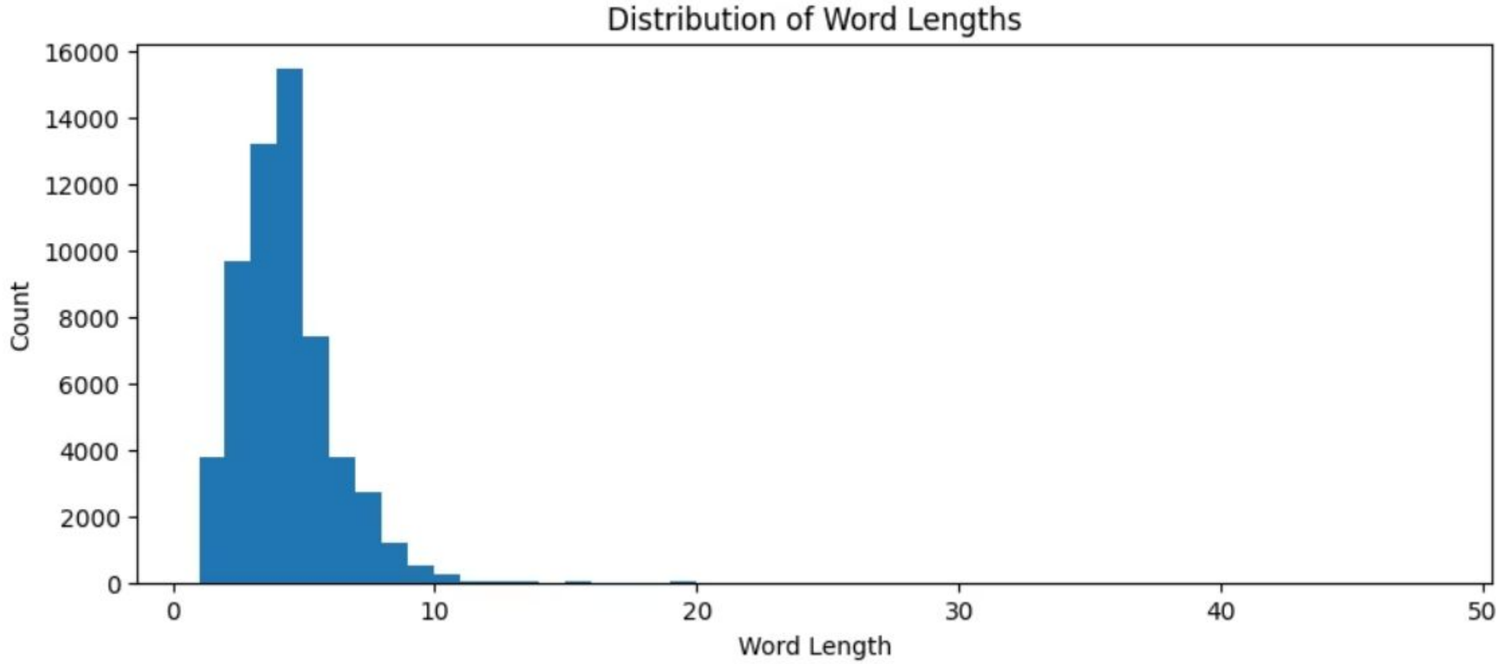
Word Statistics and Word Length Distribution

Key Metrics

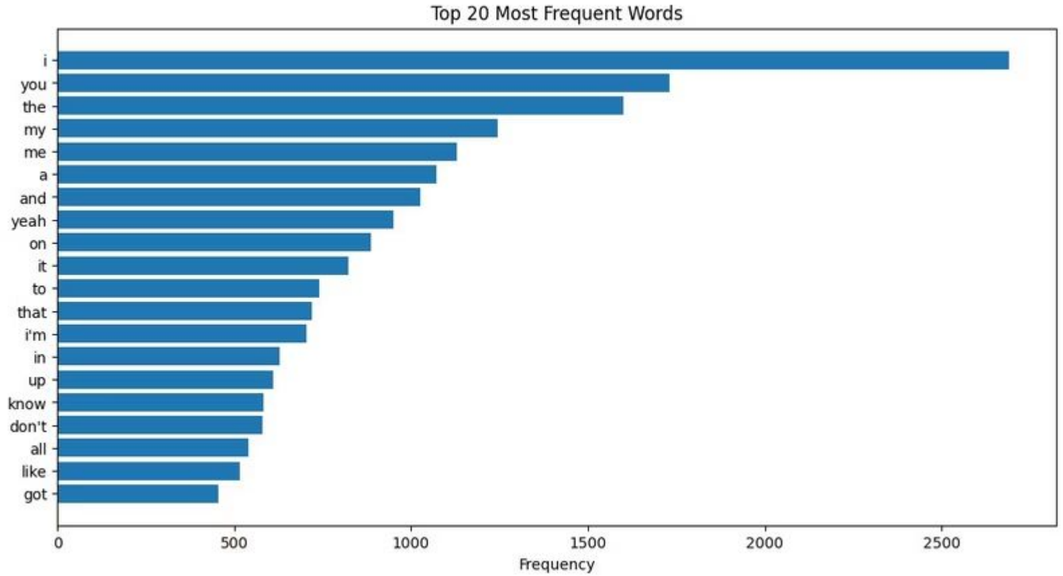
Total words in all lyrics exceed 58,000, with nearly 4,000 unique words. On average, each song contains about 393 words, reflecting rich lyrical content suitable for detailed text analysis.

Word Length Distribution

The distribution histogram shows that most words fall within short to medium-length ranges, a pattern typical in song lyrics that balances rhythm, rhyme, and clarity.



Top 20 Most Frequent Words in Lyrics



1

Common Themes and Language

The top 20 words include personal pronouns and commonly used English words such as I, you, the, and my, highlighting the personal and conversational nature of the lyrics.

2

Zipf's Law Conformance

The frequency distribution follows Zipf's expectation, affirming normative linguistic structure. This is characteristic of natural language texts including artistic song lyrics.

Contribution

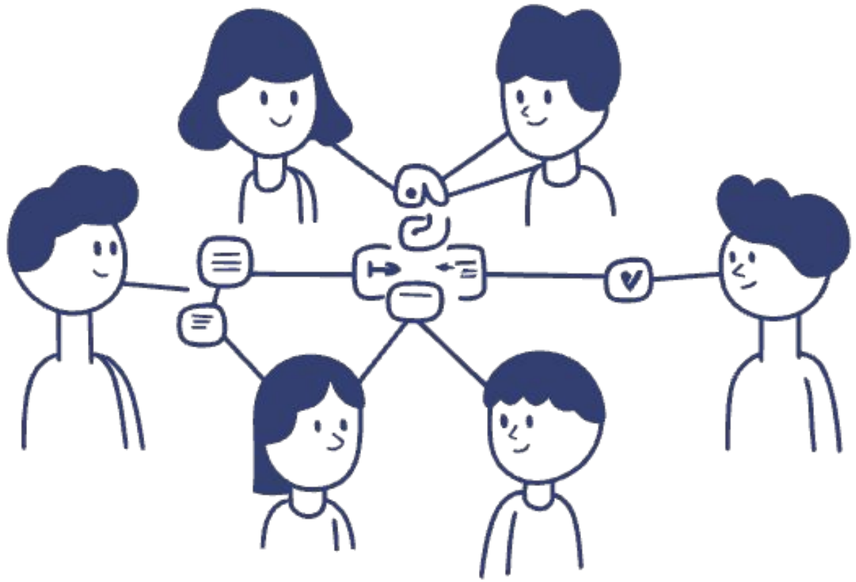
Python Notebook Creation and Dataset

Cleaning:

- Aayush Chaturvedi
- Veeramalla Saiteja

PPT Creation and Designing

- Ashvin Tiwari
- Abhinav Vinod Bhusagare
- Aayush Chaturvedi



THANK YOU!!