Necrofriggian Zipf's Law Project

This project analyzes Post Malone's lyrics using Zipf's Law.

Team Members

Aryan Verma

Coding Lead: data cleaning and Zipf’s Law implementation.

Ambuj Vashistha

Coding & Visualization: created charts and data scripts.

Ayush Kumar Singh

Research & Documentation: theoretical insights and clarity.

Nandan S Acharya

Presentation Design: final formatting and synthesis.

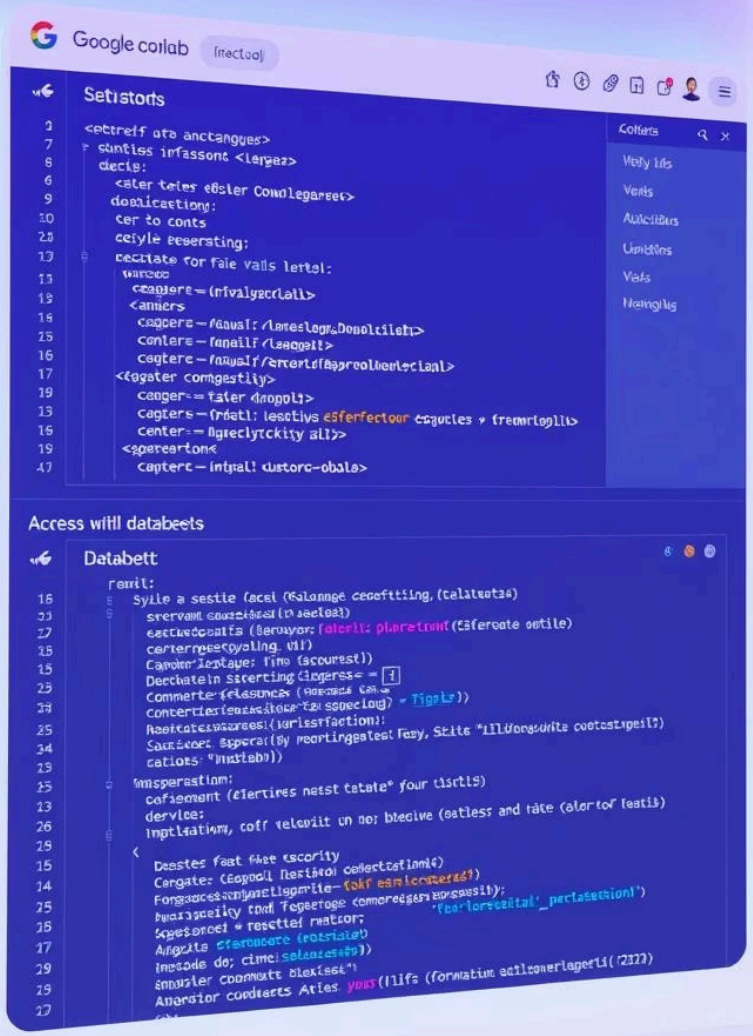
Accessing the Dataset

Dataset Source

PostMalone.csv with lyrics and metadata.

Access Method

Mounted Google Drive for direct CSV reading in Colab.



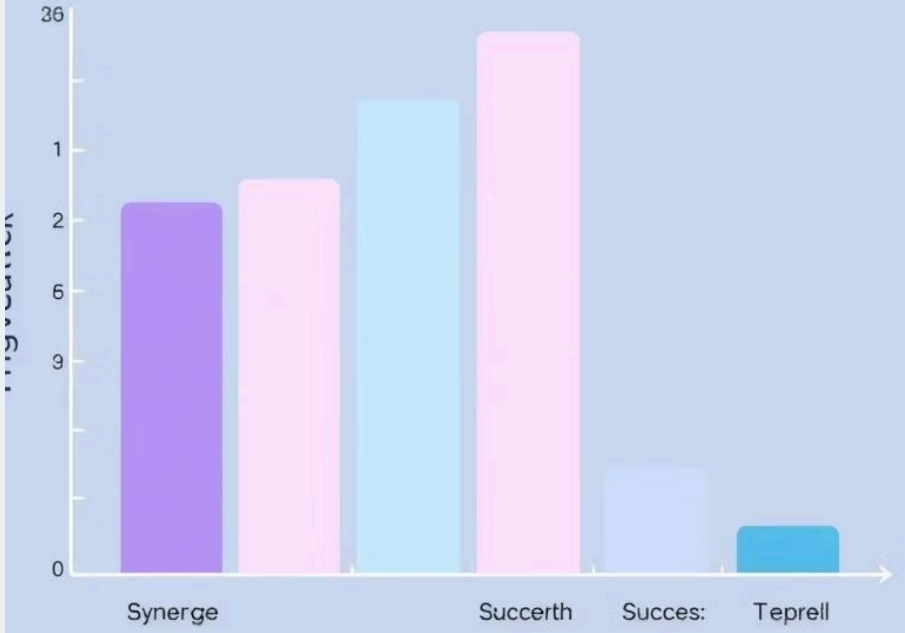
Word Frequency Analysis

Tokenization

Words counted across all lyrics.

Top Words

Extracted 30 most frequent words for Zipf's Law analysis.



Song Release Trends

Yearly Counts

Counted songs released each year.

Trend Insight

Visualized productivity and release frequency over time.

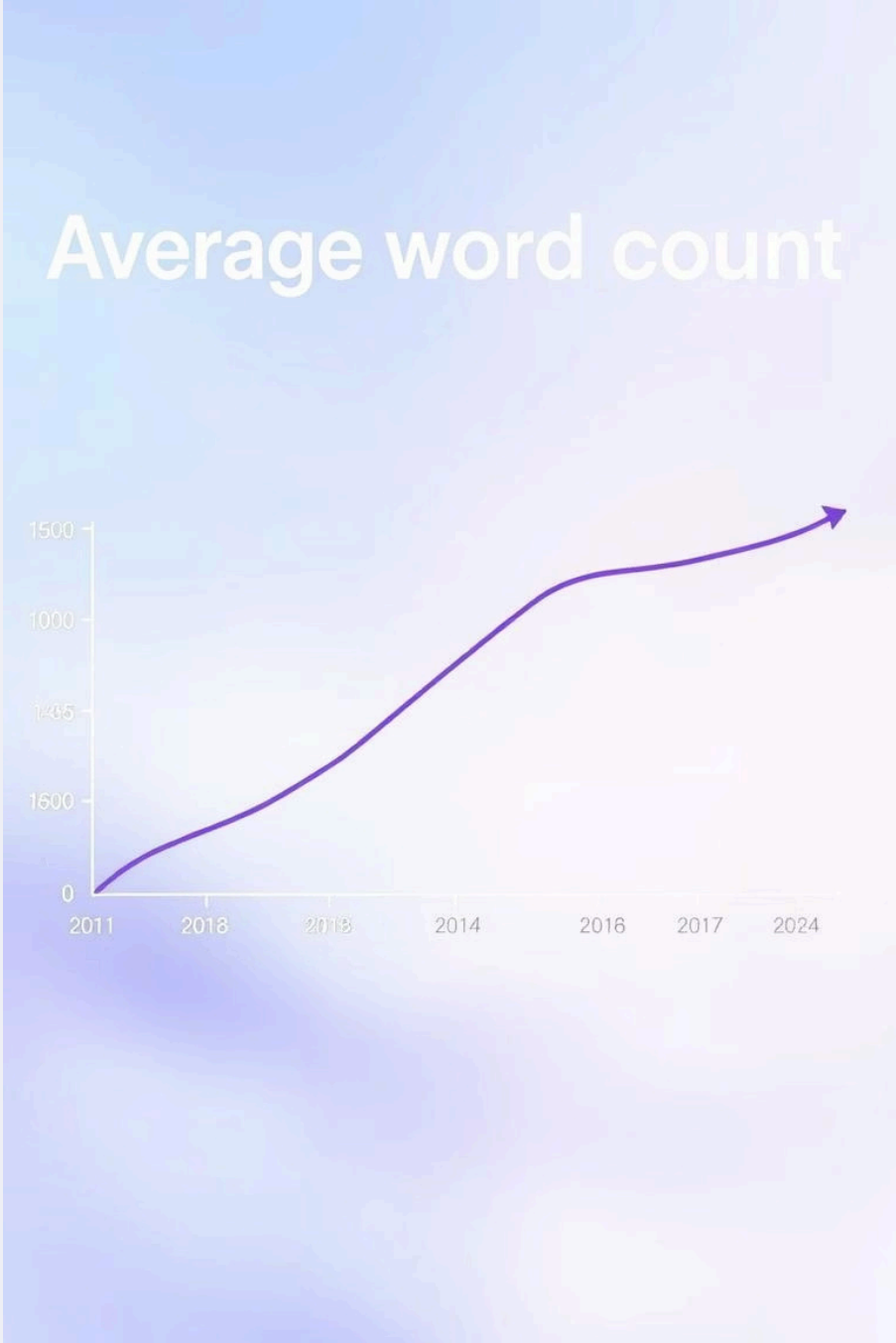
Average Word Count Over Time

Calculation

Computed word count per song and averaged by year.

Visualization

Line plot shows lyrical length changes over time.



Zipf's Law: Actual vs Expected

Comparison

Plotted actual frequencies against Zipf's theoretical model.

Observation

Word distribution mostly follows Zipf’s Law pattern.

Key Insights & Limitations

Zipf’s Law Holds

Common words rank high as expected.

Deviations Explained

Limited dataset, lyrical repetition, and style affect smoothness.

Conclusion

Lyrics follow Zipf’s Law broadly but not perfectly.