

REAL - TIME SENTIMENT ANALYSIS OF TWEETS



PRESENTED BY : SHRUTI BADAGANDI

INTRODUCTION



- Real-time sentiment analysis is a growing need in digital communication.
- Twitter is a widely used platform for opinions.
- Understanding public mood helps businesses, media, and governments.

Social media platforms, especially Twitter, are full of real-time public opinions. Analyzing these tweets can help us understand public sentiment about topics like elections, festivals, and news events. In this project, I built a system that can read, clean, translate, analyze, and visualize public sentiment from tweets written in Kannada, Hindi, and English.



DATASET OVERVIEW

The dataset used is RealTime-TweetDataset.csv, containing over 10,000 tweets in three languages: English, Hindi, and Kannada. Each tweet includes metadata such as the timestamp, user location, and raw text. The sentiment expressed in the tweet is the primary feature of interest. By leveraging this data, I can classify tweets into sentiment categories and analyze the results across different dimensions like time & Language.

TEXT PREPROCESSING & LANGUAGE DETECTION

Before analyzing sentiment, I cleaned the text to remove unnecessary parts like URLs, @mentions, and emojis. I also converted all text to lowercase and removed extra spaces. This step ensures our NLP models work on clean, normalized input, improving accuracy.

Since tweets were in different languages, I **used langdetect** to detect whether a tweet was in English, Hindi, or Kannada. Non-English tweets were then translated to English using **TextBlob**. This step made sure all tweets were in English before sentiment analysis, allowing us to use English-only models effectively.

SENTIMENT CLASSIFICATION

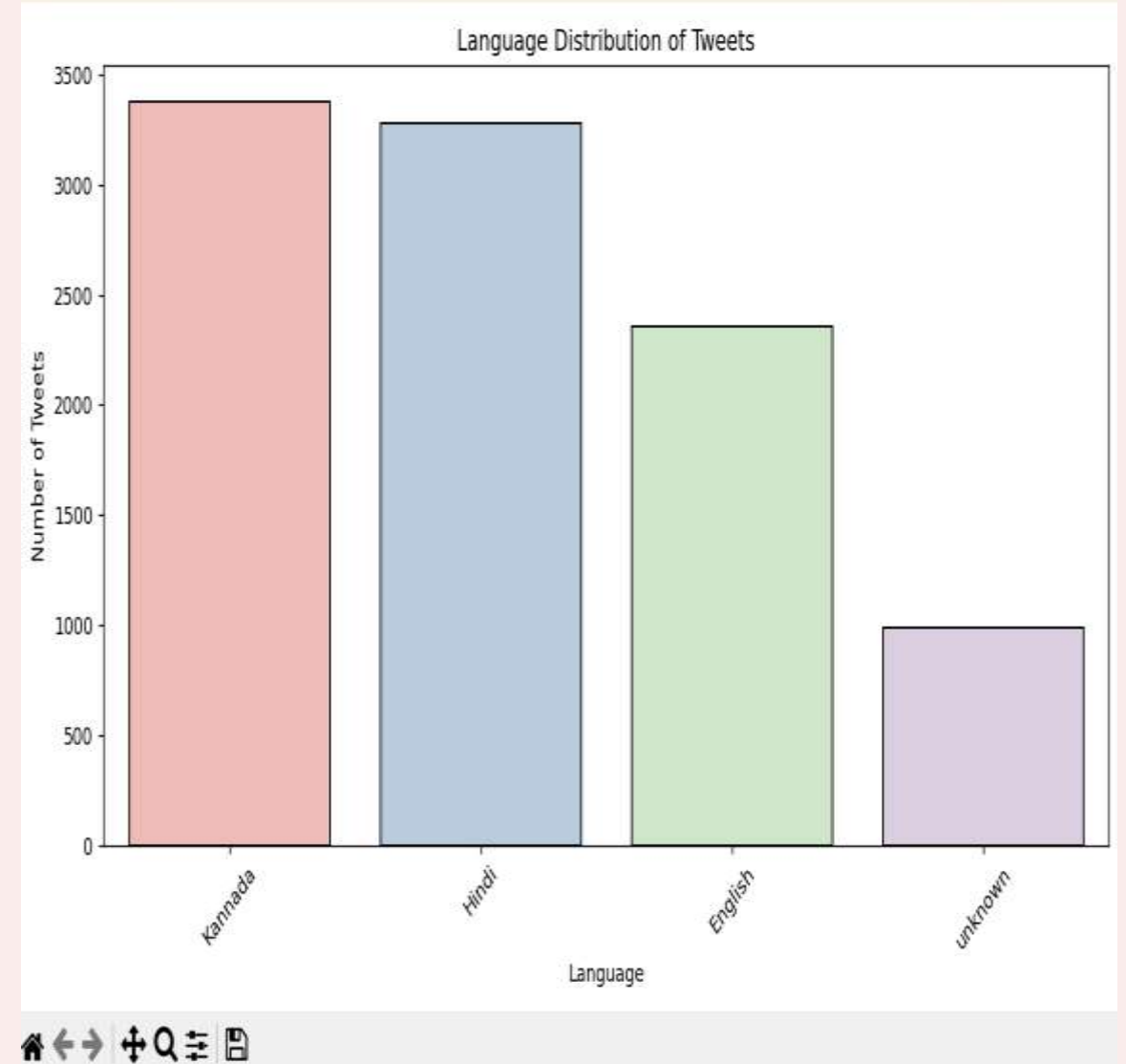
I Used the TextBlob library to assign a polarity score to each tweet. Based on polarity:

- Positive score → Positive sentiment
 - Negative score → Negative sentiment
 - Zero score → Neutral sentiment
- Each tweet was classified into one of the three categories. This method is simple yet effective and works across multiple languages when combined with pre-cleaned text.



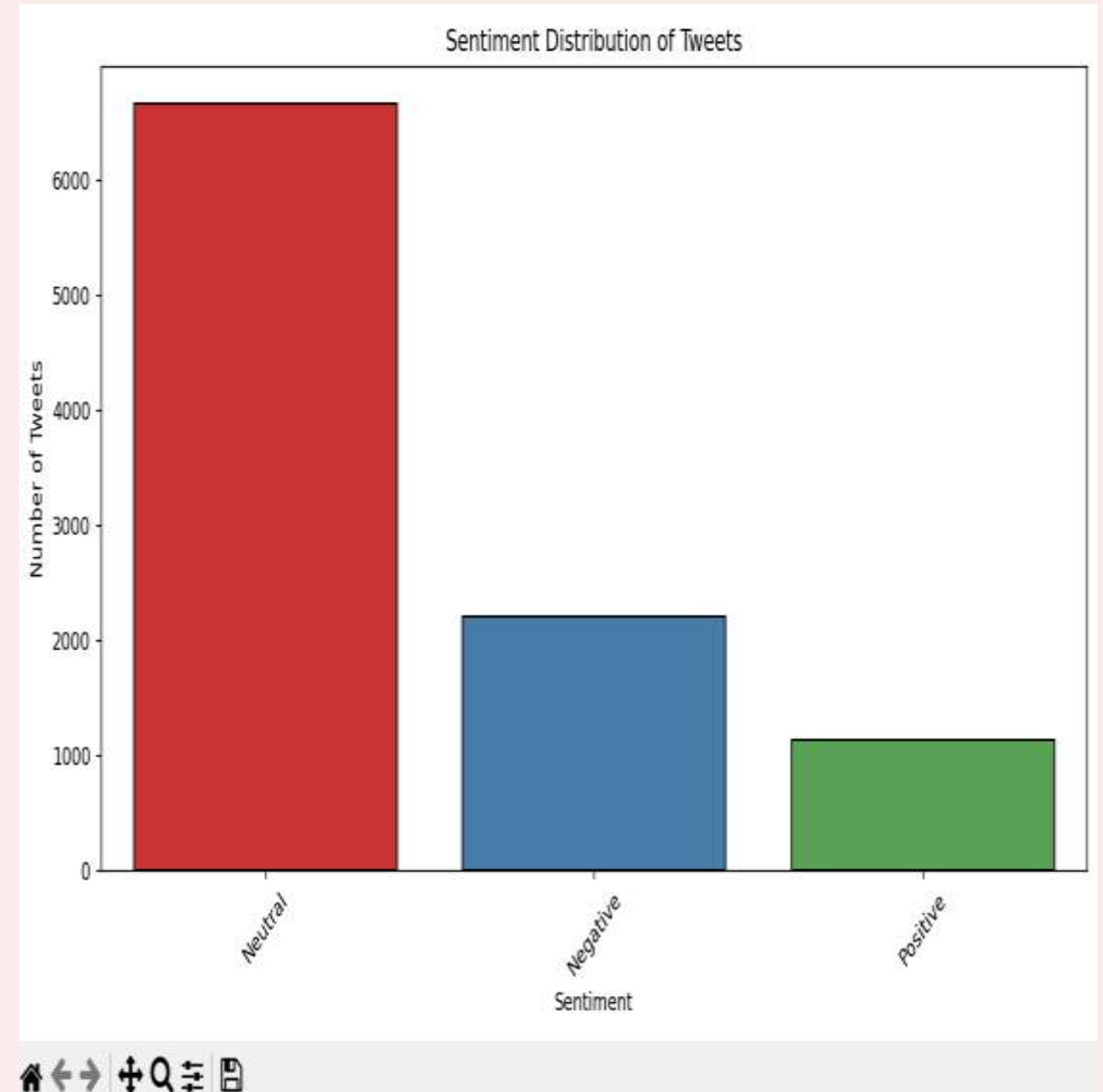
LANGUAGE DISTRIBUTION VISUALIZATION

A bar chart was used to show how many tweets were posted in each language. English tweets formed the majority, followed by Hindi and Kannada. This insight helped us understand the primary communication languages used by users and gave us direction on focusing language-specific sentiment analysis for different regions.

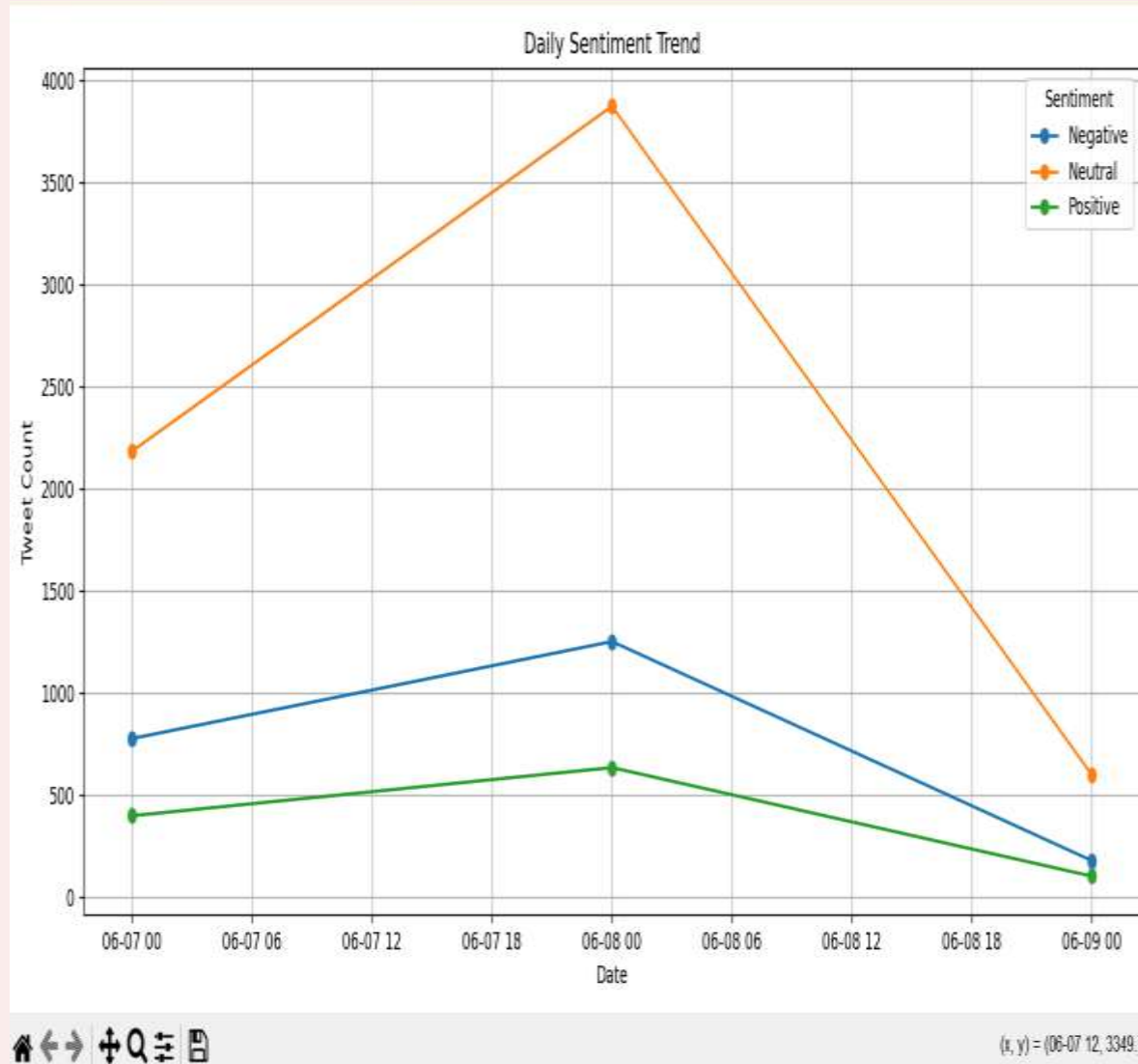


SENTIMENT DISTRIBUTION VISUALIZATION

This bar chart was used to present the distribution of tweets based on sentiment categories. We found that a large portion of the tweets expressed positive sentiment, followed by neutral and negative. This gave an overall optimistic tone to the dataset but also emphasized the need to monitor for sudden spikes in negative sentiment.



SENTIMENT TRENDS OVER TIME



Using the tweet Timestamp, we created a line graph to show how sentiment evolved on a daily basis. This helped identify trends like increases in negativity during controversial events or spikes in positivity during festivals or celebrations. Time-based analysis is vital for tracking changes in public mood.

STREAMLIT WEB APP

To make the project interactive and user-friendly, we built a web app using Streamlit. Key features:

- Upload or process real-time tweet data
 - View KPIs: total tweets, language count, location count
 - Filter by sentiment and language
 - Explore time-based and location-based sentiment charts
-
- Export filtered results to CSV The app allows users to explore data without coding, making sentiment analysis accessible to a wide audience.

Filters

Sentiment

Negative

Neutral

Positive

Language

English

Kannada

Unknown

Hindi



Real-time Sentiment Analysis of Tweets

Explore multilingual sentiment across Kannada, Hindi, and English tweets.

Total Tweets

10000

Languages

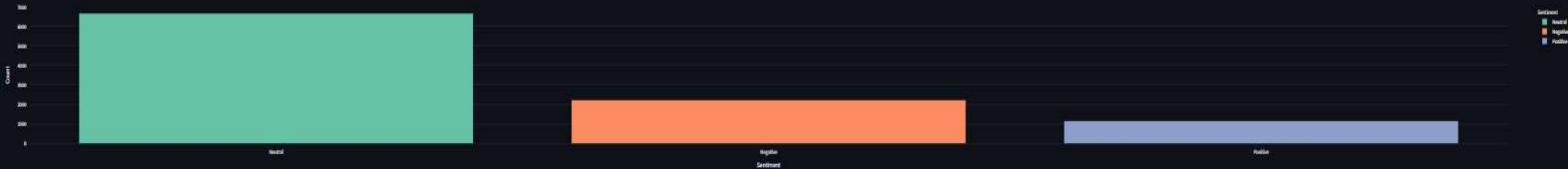
4

Locations

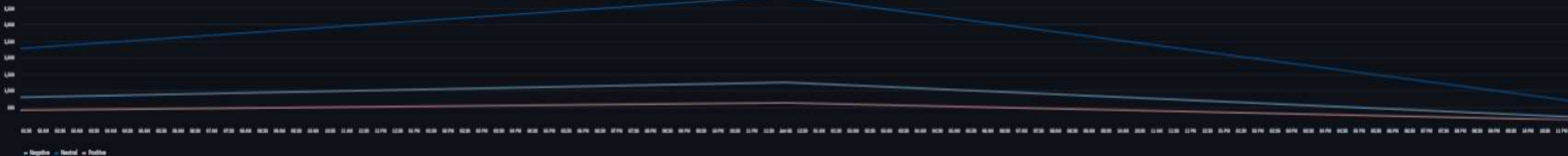
15

Sentiment Distribution

Sentiment Count

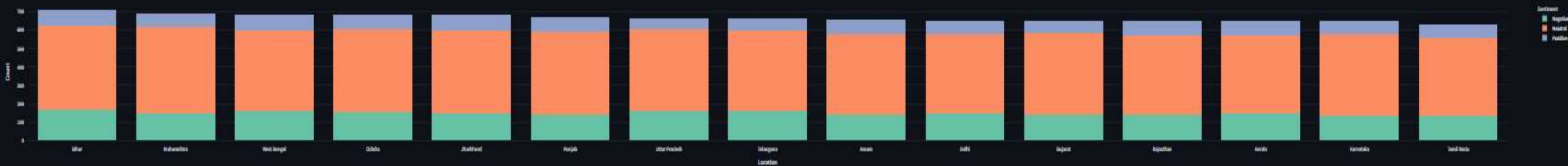


Sentiment Over Time



Sentiment by Top Locations

Sentiment in Top 15 Locations

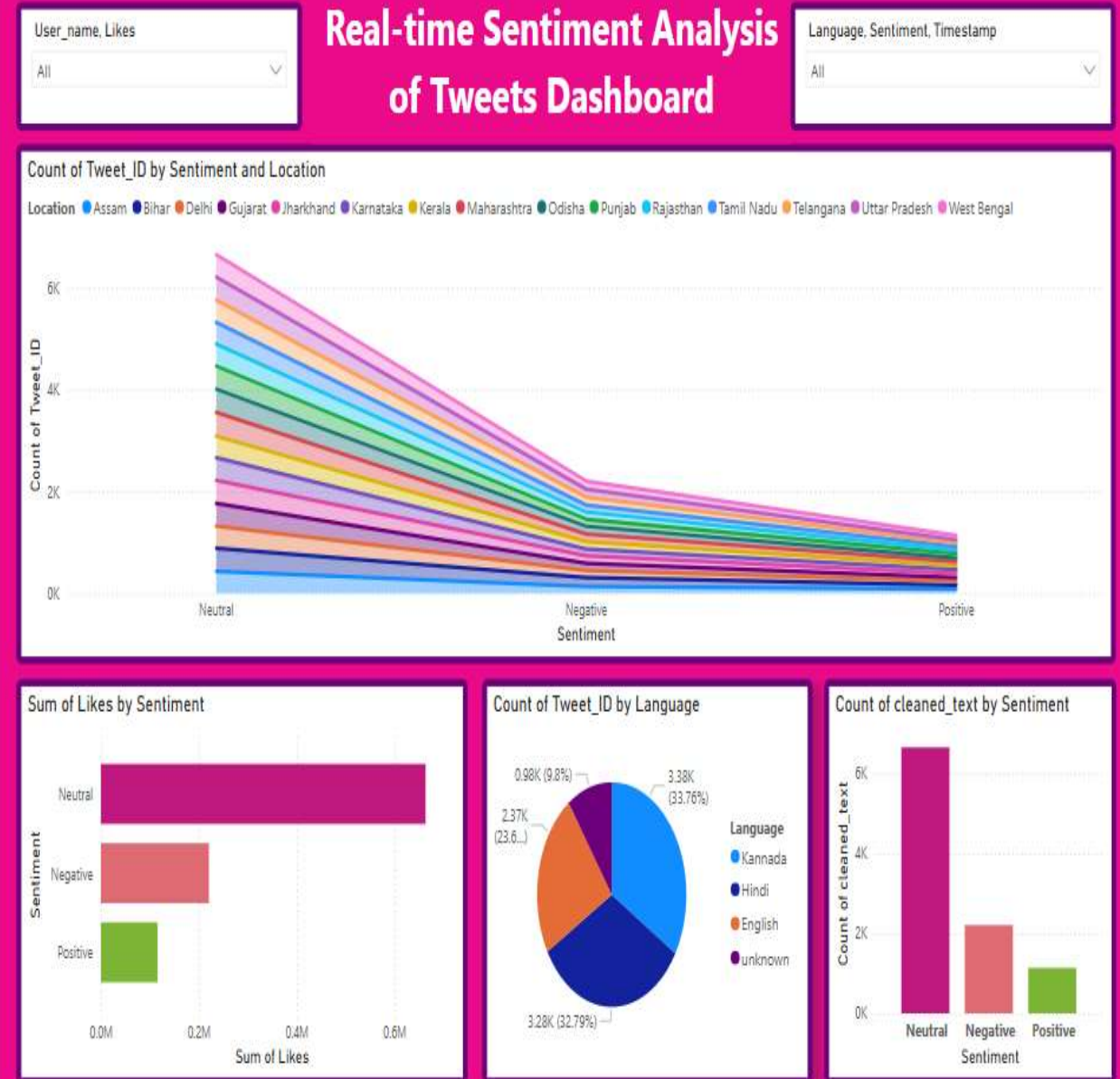


Analysis Complete Output Saved To 'CleanStream-Tweets.csv'

DASHBOARD

- Created using Power BI
- Real-time filters
- Visual insights for business users
- Maps, pie charts, bar graphs

For business audiences, we created a Power BI version of the dashboard. It allows managers to quickly review tweet sentiment using visual KPIs and maps, and it supports drill-down features for more detail.



CONCLUSION

This project showcased a full NLP pipeline using real-world social media data. From multilingual text preprocessing and language detection to sentiment classification and interactive visualizations, it demonstrates how data science can transform raw text into actionable public opinion trends. The addition of a web app makes it useful for real-time analysis in media, politics, and customer feedback.



THANK YOU