YouTube Data Analysis Project: Israel-Palestine Conflict

1. Executive Summary

Analyzed 600 videos and 167774 comments from 2014 to 2025.
Focus on sentiment evolution and stance detection (Pro-Israel vs. Pro-Palestine).

2. Methodology

Data collected via YouTube API (10-year window). Spark UDFs used for keyword-based Stance Detection.

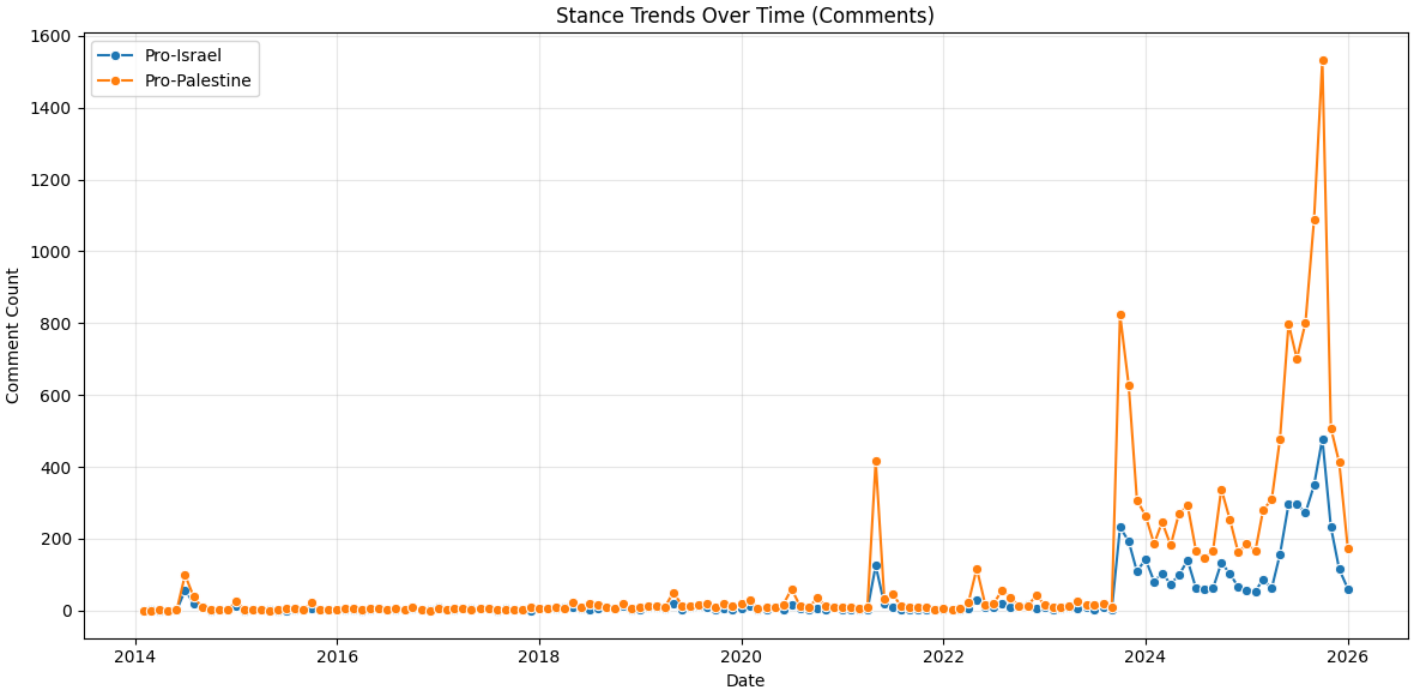
3. Data Pipeline Architecture

1. Collection: Python script -> YouTube Data API -> JSON/HDFS
2. Storage: HDFS (Raw Data)
3. Processing: PySpark (Cleaning, NLP, Aggregation)
4. Analysis: TextBlob (Sentiment), Spark ML (Topics)
5. Visualization: Matplotlib/Seaborn

4. Key Findings

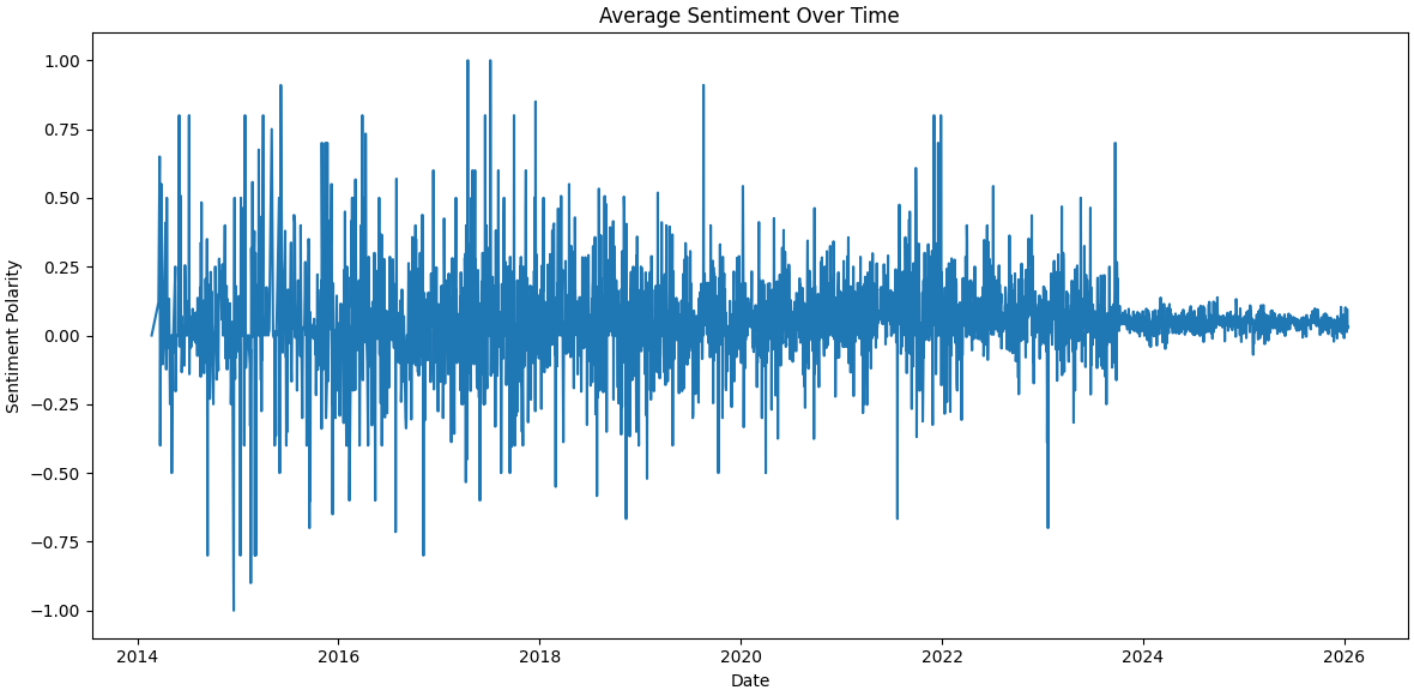
YouTube Data Analysis Project: Israel-Palestine Conflict

Stance Trend



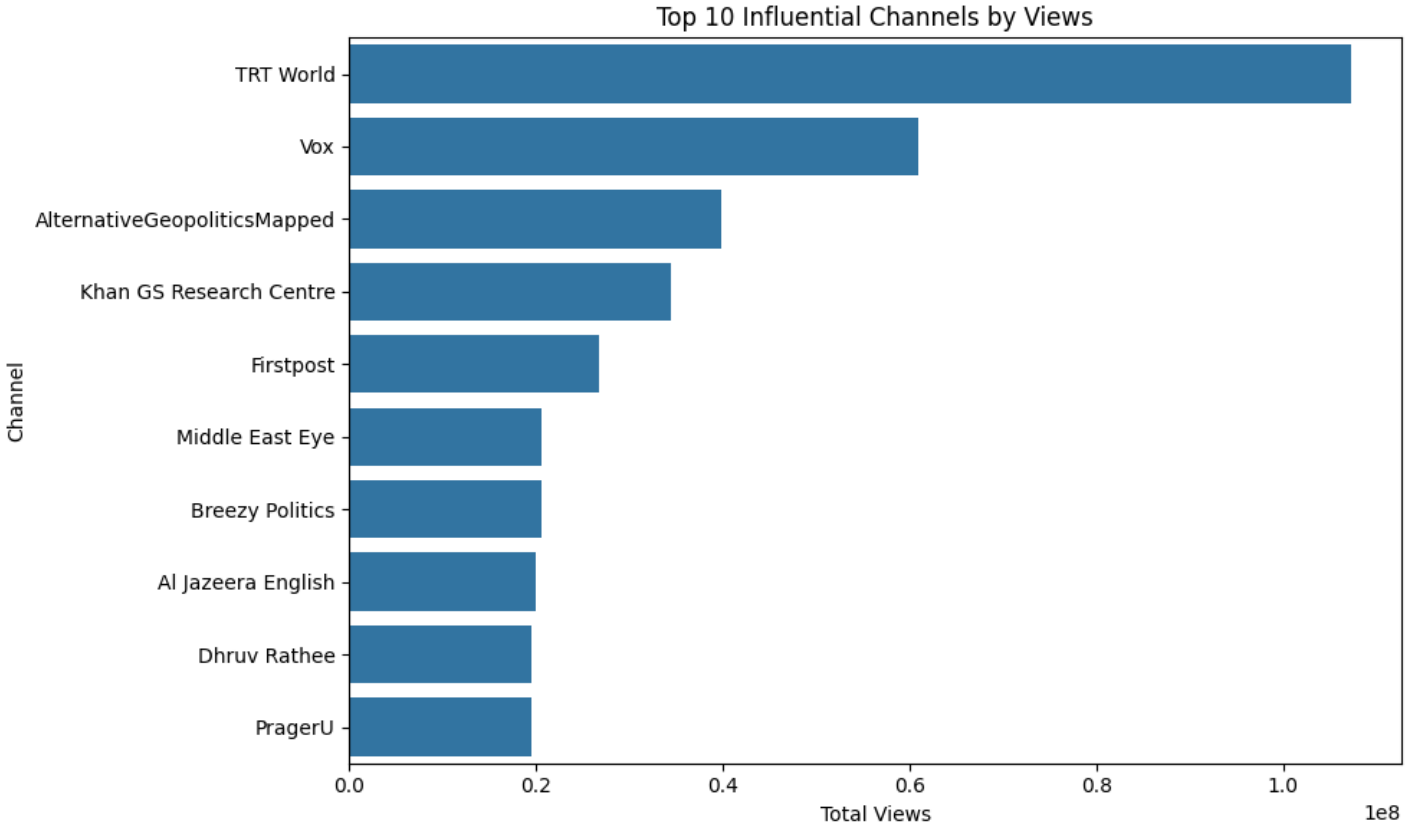
YouTube Data Analysis Project: Israel-Palestine Conflict

Sentiment Trend



YouTube Data Analysis Project: Israel-Palestine Conflict

Top Channels



YouTube Data Analysis Project: Israel-Palestine Conflict

5. Stance & Engagement Analysis

Most common stance: Neutral (148941 comments).

YouTube Data Analysis Project: Israel-Palestine Conflict

6. Conclusions & Limitations

Temporal analysis reveals correlation between conflict escalation and stance polarization.