SENTIMENT ANALYSIS

HARRIS VS. TRUMP

STEPS OF SENTIMENT ANALYSIS

1

COLLECTING THE DATA

- Gathered tweets related to Kamala Harris and Donald Trump during the debate using publicly available datasets.
- Tweets included mentions, retweets, and replies to capture a wide spectrum of public opinion.

2

CLEANING THE DATA

- Removed retweet indicators (RT), usernames (@xyz), URLs, special characters, and stop words.
- Corrected encoding issues, lemmatized the text, and identified non-English tweets with translations if needed.
- Ensured tweets were lowercased and segmented correctly to improve accuracy.

3

SELECTING SENTIMENT MODELS

Chose 4 models for comprehensive analysis:

- VADER: Specialized for short, informal text like tweets.
- BERT: Deep learning-based for contextual understanding.
- Flair: Neural network-based for sentiment prediction.
- TextBlob: Lexicon-based for general sentiment polarity.

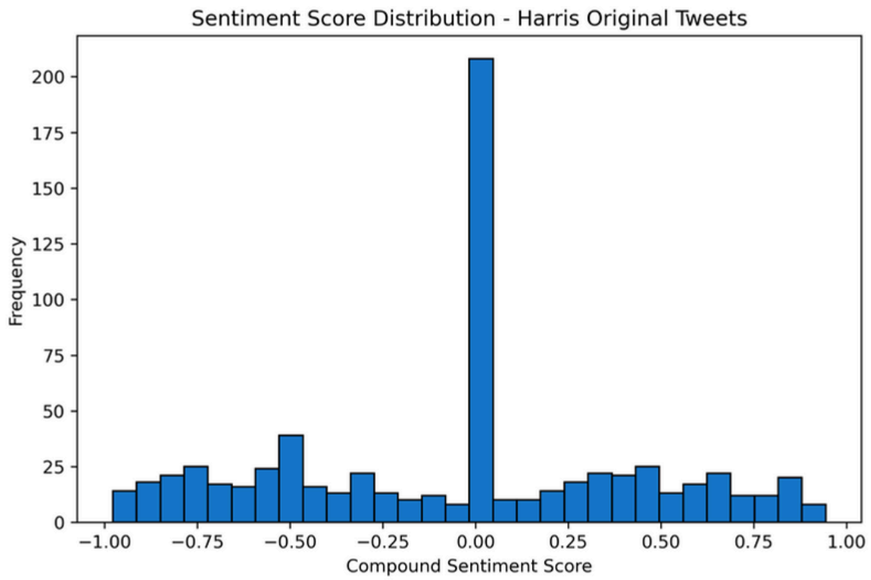
4

PERFORMING SENTIMENT ANALYSIS

- Applied the models to each cleaned tweet and assigned sentiment scores (positive, negative, or neutral). Cross-validated results across models to detect any patterns.
- Identified key sentiment trends, such as the dominance of neutral tweets and varying levels of positive/negative sentiment between Harris and Trump.

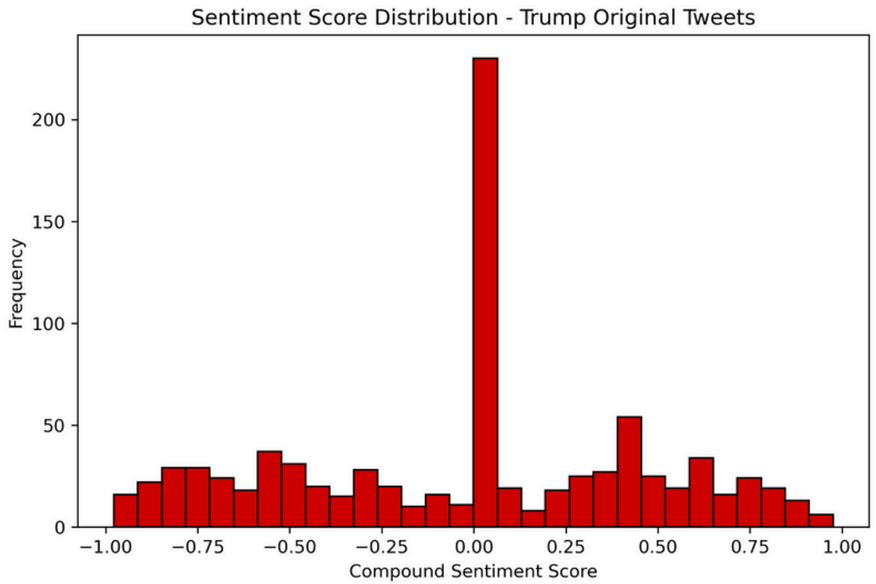
SENTIMENT ANALYSIS OF TWEETS USING VADER (KAMALA HARRIS VS. DONALD TRUMP)

Kamala Harris



- The large spike at neutral (0) indicates that many tweets were factual or without strong emotional language.
- There is a relatively balanced distribution between positive and negative tweets, suggesting mixed public sentiment.

Donald Trump



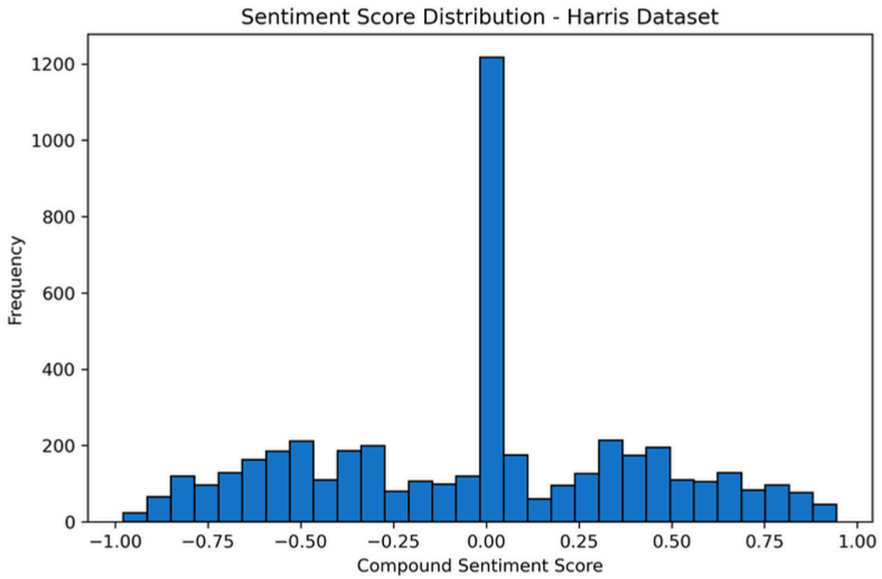
- Similar to Harris, most tweets cluster around neutral (0), but Trump's distribution skews more toward negative sentiment, suggesting more criticism or polarized reactions.
- Fewer positive tweets compared to Harris, reflecting a difference in public sentiment during the debate.

Neutral sentiment dominates both datasets, but Trump shows more negative sentiment, highlighting polarized reactions compared to Harris.

Source: X(Formerly Twitter)

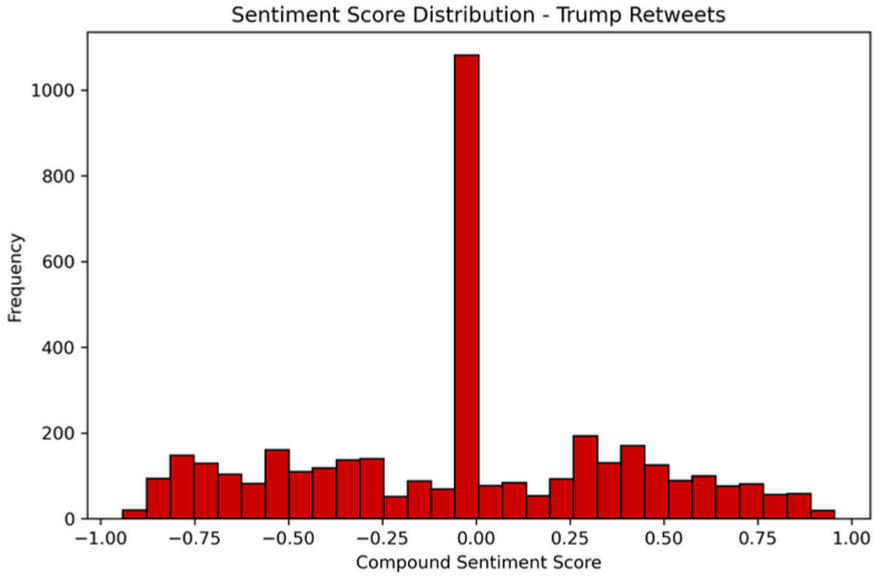
SENTIMENT ANALYSIS OF RETWEETS USING VADER (KAMALA HARRIS VS. DONALD TRUMP)

Kamala Harris



- Most retweets cluster around neutral (0), indicating the spread of factual or objective statements.
- Balanced distribution of positive and negative retweets reflects diverse public reactions being amplified.

Donald Trump



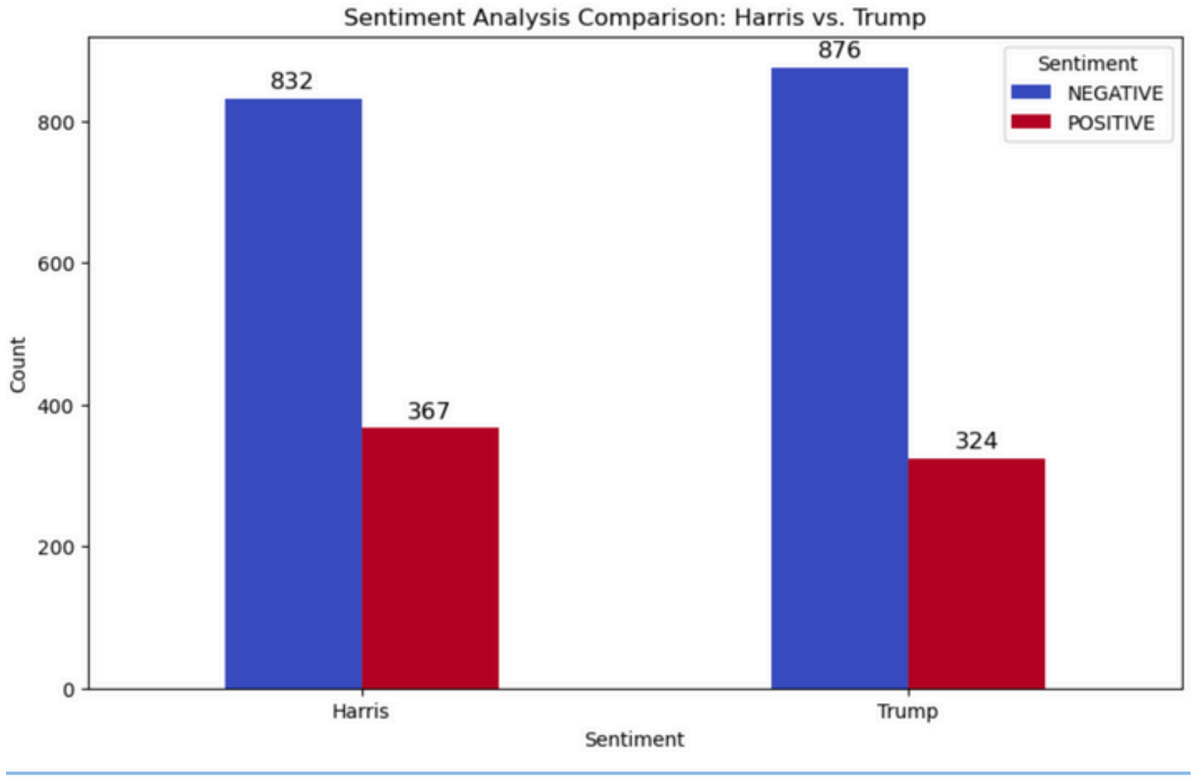
- Similar to Harris, most retweets cluster around neutral (0), but Trump's distribution skews more toward negative sentiment, suggesting more criticism or polarized reactions.
- Fewer positive retweets compared to Harris indicate less engagement with supportive content.

Retweets show a similar pattern of neutral sentiment dominance, but Trump's retweets highlight greater negative sentiment compared to Harris.

Source: X(Formerly Twitter)

SENTIMENT ANALYSIS OF TWEETS USING BERT (KAMALA HARRIS VS. DONALD TRUMP)

BERT sentiment model was designed to classify text into only Positive and Negative categories, it inherently lacks the ability to capture Neutral sentiments.



Key Observations:

- Negative sentiment dominates for both Harris and Trump.
- Trump’s tweets have slightly more negative sentiment (876 vs. 832 for Harris).
- Harris has a marginally higher count of positive sentiment tweets (367 vs. 324 for Trump).
- The overall sentiment trend suggests that both figures receive more negative engagement than positive.

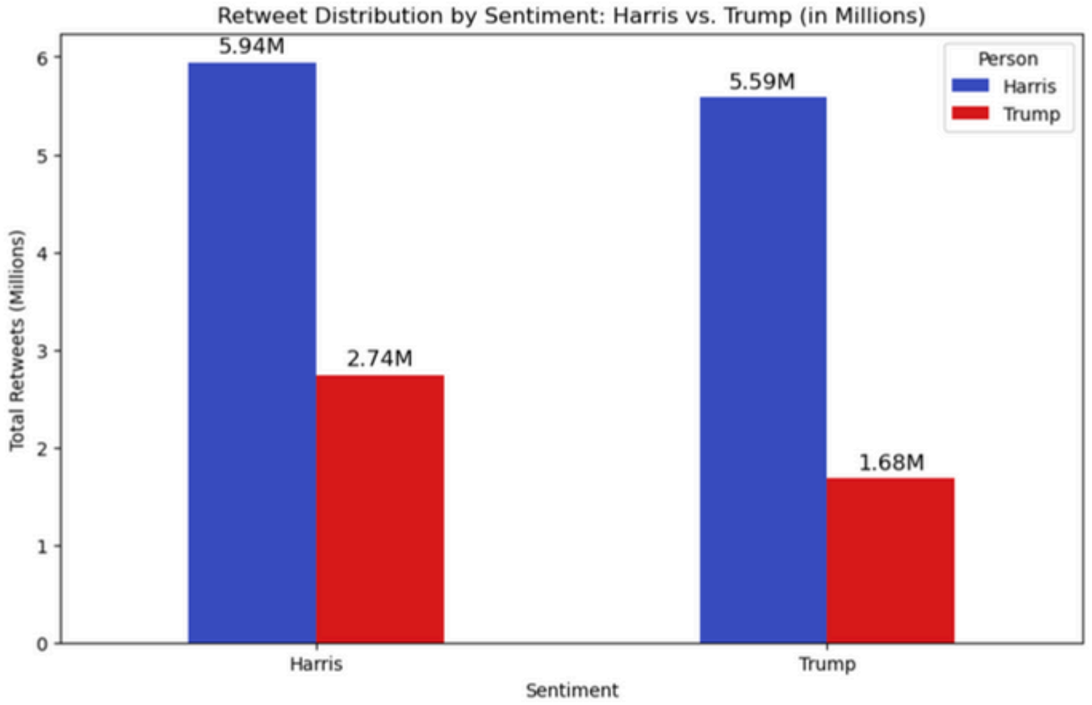
Insights & Implications:

- Negative sentiment tends to drive higher engagement, making it more widely shared.
- Trump’s tweets receive a stronger negative reaction, leading to more discussions and retweets.
- Harris’s audience engages slightly more with positive sentiment, but negative sentiment still prevails.
- The findings highlight how negative discourse is amplified online, shaping public perception and media narratives.

Source: X(Formerly Twitter)

SENTIMENT ANALYSIS OF RETWEETS USING BERT (KAMALA HARRIS VS. DONALD TRUMP)

BERT sentiment model was designed to classify text into only Positive and Negative categories, it inherently lacks the ability to capture Neutral sentiments.



Key Observations:

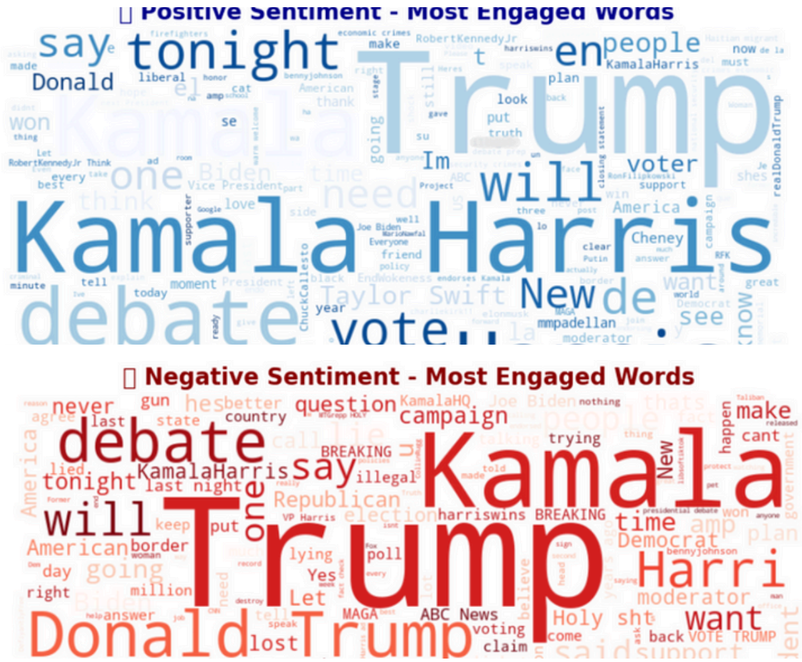
- Negative tweets receive significantly more retweets than positive ones:
  - Harris: 5.94M (Negative) vs. 2.74M (Positive)
  - Trump: 5.59M (Negative) vs. 1.68M (Positive)
- Negative tweets dominate engagement:
  - Harris: 68.46% of total retweets are from negative tweets
  - Trump: 76.84% of total retweets are from negative tweets
- Trump’s negative tweets generate more engagement than Harris’s:
  - Negative retweets: Trump (76.84%) > Harris (68.46%)
  - Trump's negative content spreads more widely than Harris’s

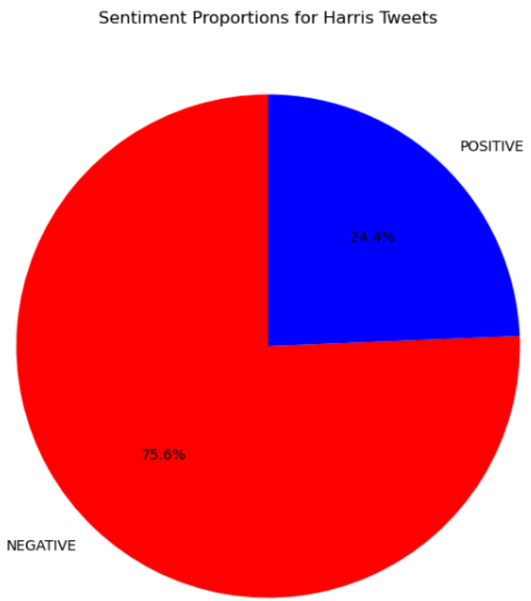
Insights & Implications:

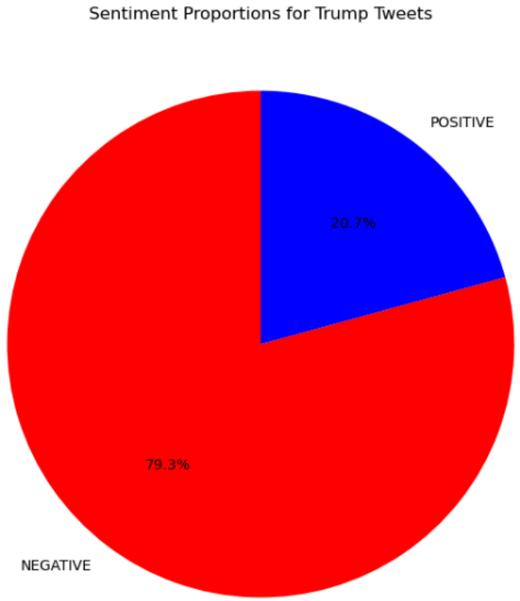
- Negative content drives higher engagement, making it more likely to be shared.
- Trump’s negative tweets gain more traction than Harris’s, showing a stronger audience reaction.
- Harris’s audience engages more with positive content compared to Trump’s, but negative sentiment still dominates.
- Negative political discourse gets amplified, influencing public perception and online discussions.

Source: X(Formerly Twitter)

SENTIMENT ANALYSIS OF VIRAL TWEETS | INSIGHTS WITH FLAIR





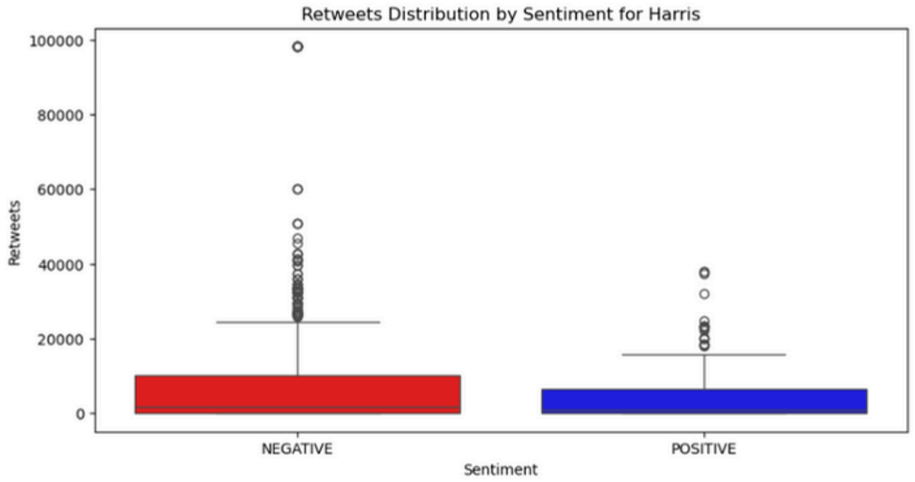


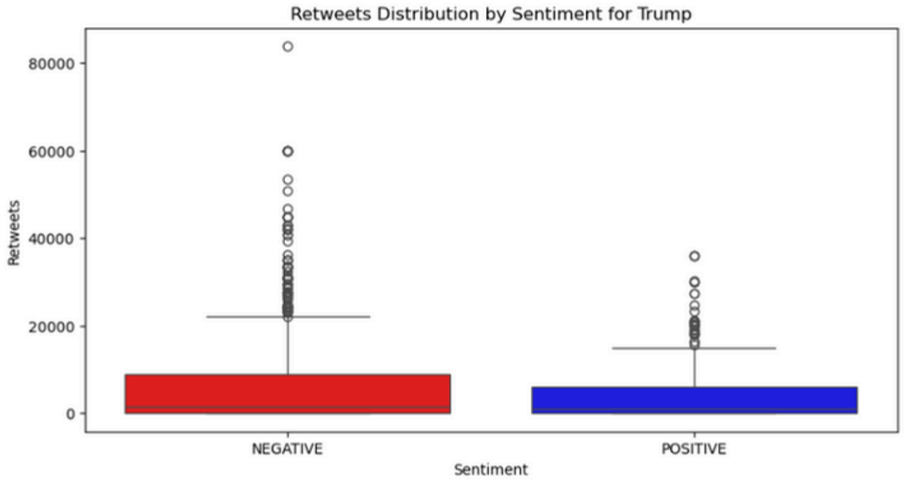
OVERALL DISTRIBUTION OF SENTIMENTS

- Trump’s tweets tend to be a little more negative (**79.3%**): Compared to Harris's (**75.6%**), Trump's tweets tend to carry a higher negative sentiment reflecting that much more critical or contentious tone.
- Harris' tweets showed more positive engagement: Harris's tweets had **24.4% positive** sentiment, more than Trump's **20.7%,** hence indicating a relatively more optimistic messaging style.
- Both of them face colossal criticism: with over three-quarters of all tweets being negative, public discourse about both leaders seems to be dominated by criticism or controversy.
- The difference is **small** yet **notable**: though Trump's tweets are a bit more negative in nature, Harris does have a marginally better public sentiment balance in this dataset.

Source: X(Formerly Twitter)

ENGAGEMENT & SENTIMENT TRENDS



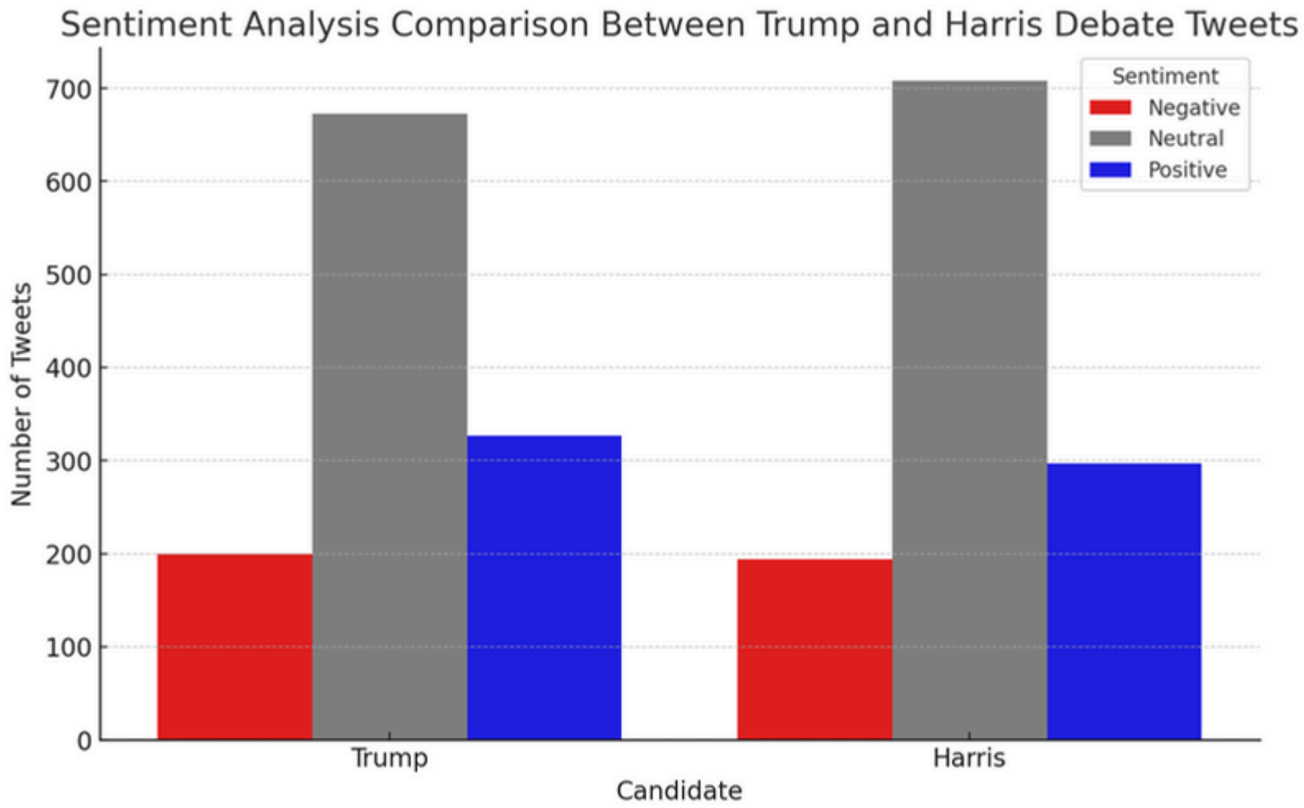


- For both Harris and Trump, **negative** sentiment tweets have a **higher retweet** count compared to positive ones, which may indicate that negative content is more engaging.
- There are several **extreme outliers** for the **negative** category in both leaders meaning that highly viral tweets (with >80,000 retweets) tend to be negative.
- There is less dispersion of the distribution in positive sentiment so there are fewer outliers in that category. In this respect, positive tweets have a lesser amount of spreading than others.
- The standard deviation for Trump is a bit higher than Harris, meaning though both get criticized, his negative tweets will receive more varying responses. **The pattern in both leaders goes like this**: The overall trend for both indicates that negative tweets tend to spread more therefore reinforcing the concept that controversial or critical content moves faster on social media.
- **Flair struggles** with **context-dependent words** appearing in both positive and negative sentiments, making it less reliable for nuanced classification.It lacks built-in stopword removal and URL filtering, requiring extra preprocessing.

Source: X(Formerly Twitter)

SENTIMENT ANALYSIS USING TEXTBLOB (KAMALA HARRIS VS. DONALD TRUMP)

Kamala Harris Vs. Donald Trump



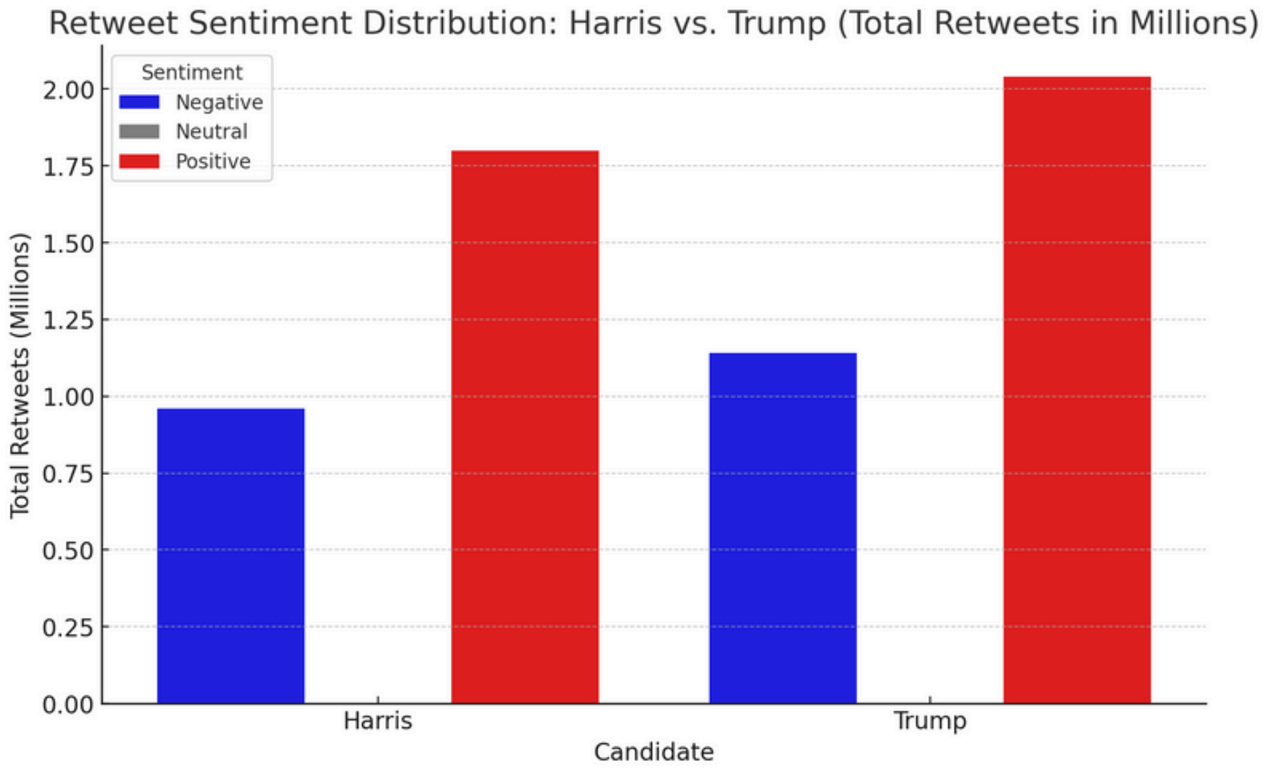
KEY INSIGHTS:

- Neutral sentiment dominates for both Trump and Harris, indicating that most tweets are factual, commentary-based, or lacking strong emotional expressions.
- Positive sentiment is slightly higher than negative sentiment for Harris, while Trump's sentiment distribution skews slightly more negative. This suggests that Harris's tweets had a more balanced or favorable engagement overall.
- Trump's sentiment distribution shows slightly more negative tweets compared to Harris, suggesting that there was a slightly stronger pushback or criticism toward him in this dataset.

Source: X(Formerly Twitter)

SENTIMENT ANALYSIS USING TEXTBLOB (KAMALA HARRIS VS. DONALD TRUMP)

Kamala Harris Vs. Donald Trump Retweets



KEY INSIGHTS FROM THIS VISUALIZATION:

- Trump’s total retweets are higher than Harris’s, with over 2 million positive retweets, 1.1 million negative retweets, and 579 neutral retweets, indicating widespread engagement with both factual content and opinion-driven tweets.
- Harris’s total retweets are slightly lower, with 1.8 million positive retweets, around 960K negative retweets, and 623 neutral retweets, suggesting that many shared discussions were commentary-based or factual rather than strongly opinionated.
- Positive tweets receive more retweets than negative ones for both candidates, indicating stronger engagement with supportive content.
- Trump’s negative tweets still receive significant engagement, suggesting that critical content about him spreads widely.
- Neutral retweets contribute to overall engagement, with Harris’s neutral retweets slightly outpacing Trump’s (623 vs. 579), reflecting the role of fact-based or less emotionally charged discussions in shaping public discourse.

Source: X(Formerly Twitter)

THANK YOU