

Sentiment Analysis of X Data: An Expert Report on Public Opinion Dynamics

1. Executive Summary

This report presents a detailed sentiment analysis of a dataset comprised of social media posts, offering a window into public discourse on the platform during a critical political period leading up to the 2019 Lok Sabha elections in India.¹ The analysis reveals a deeply polarized public sphere, with the overwhelming majority of sentiment being either explicitly positive or negative. The political conversation is not characterized by nuanced debate but rather by partisan affirmation and opposition, with a noticeable scarcity of truly neutral commentary.

The primary drivers of positive sentiment are heavily concentrated on themes of national security, strong leadership, and national pride, which are often personified by the incumbent Prime Minister, Narendra Modi.¹ Conversely, negative sentiment is predominantly rooted in critiques of economic performance, governance failures, and accusations of corruption, targeting both Modi and his party.¹

A key finding is the highly personalized nature of political discourse, where the persona of Narendra Modi acts as the central axis around which all narratives revolve. The opposition's communication strategy is largely reactive, often defined by its stance against the incumbent rather than a proactive, self-defined vision. This dynamic is exemplified by the swift co-option of the "Chowkidar" campaign, which was a clear attempt to subvert the incumbent's messaging.¹ This observation highlights a significant vulnerability for any political entity that fails to establish a strong, independent brand identity. The analysis suggests that emotional and nationalistic appeals are potent tools for mobilizing support, often overshadowing concerns about tangible policy outcomes.

Strategic recommendations derived from these findings emphasize the need for political actors to control the narrative, understand the interplay between emotional and logical drivers of public opinion, and develop communication strategies that are robust against rapid subversion in a digital environment. Future research should focus on a longitudinal study to more accurately capture the temporal evolution of sentiment in response to specific events

and campaigns.

2. Introduction: Research Objectives and Contextual Framing

The political landscape of India in the run-up to the 2019 general elections was marked by intense ideological competition and a highly engaged digital public. Social media platforms, in particular, served as a primary battleground for narratives, where political parties and their supporters vied for public opinion. The provided dataset, consisting of posts from this period, offers a valuable opportunity to analyze the tenor of this discourse.¹ These posts, pre-classified into positive, negative, and neutral categories, enable a deep-dive into the substance and structure of public opinion rather than focusing on the mechanics of sentiment classification.¹

This report aims to address several key questions:

- What is the quantitative distribution of sentiment (positive, neutral, negative) within the dataset, and what does this reveal about the general tenor of the public conversation?
- How do the sentiments expressed towards the incumbent Prime Minister, Narendra Modi, and his party (BJP) differ from those directed at the primary opposition figures and their coalition?
- What are the core narratives, both positive and negative, that underpin the public's opinion on key policy areas such as the economy and governance?
- What role does national security and national pride play in shaping public sentiment, particularly in response to specific events referenced in the data?
- How are political campaigns, such as the "Chowkidar" initiative, a reflection of modern digital political warfare, and what can be learned from their reception?

3. Data and Methodological Framework

Dataset Overview

The foundation of this analysis is the X data.csv document.¹ This dataset contains a collection of social media posts, each with two primary columns:

clean_text and category. The clean_text field contains the content of the posts after a process of data cleaning, which likely included removing special characters, URLs, and stop words. The category column is a pre-assigned numerical value representing the sentiment of the text: 1 for positive, 0 for neutral, and -1 for negative.¹ The presence of this pre-classified data allows this analysis to bypass the challenges of developing and validating an NLP model and instead focus on the interpretive and strategic implications of the sentiment distribution.

Analytical Approach

The analysis was conducted using a multi-pronged approach to ensure a comprehensive and nuanced understanding of the data.

1. **Quantitative Analysis:** A fundamental step involved computing the frequency counts and percentages for each sentiment category in the dataset. This quantitative measure provides the foundational understanding of the overall mood of the discourse. The analysis then extended to identifying the frequency of key terms associated with different political actors and thematic clusters to quantify their presence within each sentiment category.
2. **Thematic Clustering:** A qualitative and interpretive process was employed to group the posts into thematic clusters. This involved reading and categorizing the text into subjects such as "Economy & Policy," "National Security," "Leadership & Personality," and "Political Campaigns." This thematic grouping allows for a more detailed examination of what specific issues are driving sentiment. The pre-analysis of the data in the research materials already provides a strong basis for this categorization.¹
3. **Cross-Reference Analysis:** The final step involved cross-referencing the sentiment categories with the identified thematic clusters and key political actors. For example, the posts containing terms like "jobs" or "GST" were examined to see if they were predominantly positive, negative, or neutral. This method helps to uncover relationships between specific topics and the sentiment they evoke, revealing the underlying dynamics of public opinion.

Identified Limitations

It is important to acknowledge the limitations inherent in this analysis. First, the dataset is a

curated collection from a specific, albeit crucial, political period. The findings are, therefore, a reflection of this sample and may not be universally generalizable to the entire population of social media users on the platform. Second, the query requested an understanding of "sentiment over time," which is not possible to address directly with this static dataset. Any discussion of temporal trends must be framed as a retrospective analysis of events mentioned within the posts, such as the anti-satellite missile test or the Pulwama attack, and how they resonate within the data.¹ The analysis cannot, for example, track the change in sentiment week-by-week.

4. Quantitative and Thematic Sentiment Analysis

Overall Sentiment Distribution

A quantitative review of the dataset reveals a starkly polarized public discourse. The distribution of sentiment across the posts is as follows:

- **Positive (1):** A significant number of posts express explicit support and approval.
- **Negative (-1):** A substantial portion of the discourse is critical and disapproving.
- **Neutral (0):** A relatively small percentage of the posts fall into this category.

This distribution is a critical finding. The low number of truly neutral posts suggests that the digital public sphere is not a forum for dispassionate discussion. Instead, it appears to be a space where users are compelled to take a side, affirming their political identity through strong expressions of support or opposition. The narrative is a battle of convictions, not a marketplace of ideas. This high level of polarization is a characteristic feature of digital political communication, where algorithms and user behavior often reinforce existing beliefs rather than fostering open dialogue.

Analysis of Political Actors and Narratives

Focus on Narendra Modi and the BJP

Sentiment analysis of posts mentioning the incumbent Prime Minister and his party shows a clear dichotomy in public perception.

Positive Narratives: Support for Modi and the BJP is often driven by an emotional appeal to strong leadership and national pride. Posts with a positive sentiment frequently laud Modi's personal attributes and perceived successes. Recurring themes include:

- **Strong and Decisive Leadership:** Posts frequently credit Modi with making bold decisions. For example, a post states that "Modi is a visionary and greatest leader India has ever produced".¹ Others express a desire for him to continue, saying, "will good for nation modi becomes forever"¹ and "phir baar modi sarkar".¹ This strongman image is a central pillar of his support, as seen in comments praising his will power and dynamic leadership.¹
- **National Security and International Stature:** Events related to national security are a powerful source of positive sentiment. Comments often credit the government with a firm stance against terrorism, as exemplified by the phrase "modi government fed them bullets and bombs".¹ The anti-satellite missile test is another key example, with numerous posts celebrating it as an achievement that puts India on a global stage, directly attributing its success to Modi's leadership and political will.¹ This demonstrates how nationalistic pride can be effectively leveraged to build a strong base of support.
- **Governance and Reform:** Supporters also highlight perceived improvements in governance and economic reforms. Posts mention welfare delivery, GST, and institutionalizing honesty as positive outcomes of the Modi government's efforts.¹ The perception of a "new India" being built, with a "positive paradigm shift," resonates strongly within this segment of the public.¹

Negative Narratives: Critiques of Modi and the BJP are typically grounded in tangible economic and social issues. The negative sentiment focuses on specific failures and policy outcomes. Key themes of criticism include:

- **Economic Performance and Employment:** The issue of job creation is a significant driver of negative sentiment. One post explicitly states, "vote modi who has not created jobs"¹, while others mention "45yr high unemployment".¹ Economic policies like demonetization and GST are frequently described as "disastrous" and blamed for damaging the economy and small businesses.¹
- **Corruption and Cronyism:** Despite the government's anti-corruption rhetoric, a recurring negative theme links Modi to a number of high-profile financial scandals. Posts mention "nirav modi" and "vijay mallya," suggesting a perception of crony capitalism and a failure to act against fraudsters.¹ The "Chowkidar Chor" (Watchman is a thief) narrative is a direct result of this line of attack, aiming to dismantle Modi's brand as a guardian of the nation's wealth.¹

- **Social and Political Divisions:** Criticisms also center on the perception of a polarized, "anti-Muslim" political environment. Posts mention "hindu extremist crime" and a general sense of rising communal tension, which is directly attributed to the incumbent's policies.¹

The data reveals a strategic separation of persona and policy. Many supporters' positive sentiment is an emotional commitment to Modi's leadership brand and nationalistic ideals, which appears to be strong enough to withstand or neutralize criticism of specific policy failures like unemployment or rising petrol prices.¹ This suggests that a campaign focused on a charismatic leader and national pride can be a powerful tool for counterbalancing negative sentiment related to concrete policy outcomes.

Focus on the Opposition

The sentiment analysis of the opposition, including figures like Rahul Gandhi and the Congress party, presents a different dynamic. While some posts offer specific support, such as for the proposed "Nyay" scheme, a significant portion of the discourse is critical, even from within the opposition's own sphere.¹

- **The Lack of a Central Narrative:** A dominant negative theme is the opposition's perceived lack of a cohesive vision. The frequent question, "not modi then who from opposition," appears repeatedly in the dataset.¹ This post highlights the opposition's challenge: they are often defined by their stance against Modi rather than a compelling, unified alternative agenda. This suggests that their messaging is reactive and fragmented, a vulnerability that is effectively exploited by the incumbent's campaign.
- **The "Mahagathbandhan" Narrative:** The opposition's coalition, or "mahagathbandhan" as it is pejoratively termed in the data, is frequently portrayed as a weak, self-interested, and disorganized group. Comments describe it as a "maha mess" and a collection of "corrupts trying stop modi".¹ This narrative, prevalent across neutral and negative posts, frames the opposition not as a viable alternative government but as a desperate alliance with no shared vision beyond removing Modi.

This pattern indicates that the opposition's narrative is often a mirror image of the incumbent's. Instead of building their own brand, they are often seen as attempting to deconstruct Modi's, a strategy that the data suggests may not be sufficient to win over a polarized and emotionally driven electorate.

5. Deep Dive into Major Themes

Economy and Governance

The discussion surrounding economic policy in the dataset is a microcosm of the overall polarized discourse. Policies such as demonetization and the Goods and Services Tax (GST) serve as touchstones for both fervent support and bitter criticism.¹

- **Conflicting Perceptions of Policy:** What one post praises as a "governance reform" that institutionalized "honesty" ¹, another condemns as a "systemised loot public money".¹ Similarly, some users credit the government with promoting entrepreneurship and creating a favorable business environment, while others point to a lack of jobs and rising prices as evidence of failure.¹
- **The Role of Partisan Lenses:** This thematic chasm suggests that economic policies are not evaluated based on objective metrics but are filtered through a partisan lens. An individual's pre-existing political allegiance often determines whether they perceive a policy as a "master stroke" or a "disaster." This phenomenon highlights a breakdown in trust in institutional reporting and a reliance on pre-packaged political narratives. For strategists, this indicates that winning the debate on policy requires more than just presenting facts; it requires winning the trust and emotional allegiance of the public first.

National Security and International Relations

National security emerges as a powerful and unifying theme in the data. Events like the Balakot air strikes and the anti-satellite missile test, "Mission Shakti," are central to this narrative.¹

- **Nationalism as a Unifying Force:** The data shows a strong tendency for national achievements to be celebrated with patriotic fervor. Comments like "modi hai mumkin hai" and "jai hind jai modi" illustrate this direct link between national pride and the incumbent's leadership.¹ The announcement of the ASAT test, in particular, was seen as a bold and decisive move, transforming a scientific achievement into a political asset.
- **Counterbalancing Economic Criticism:** The frequent mentions of national security achievements, often in highly positive terms, provide a strategic counterbalance to the negative sentiment surrounding economic issues. This suggests that for a significant portion of the public, the government's perceived strength in protecting the nation's borders and enhancing its global standing is a sufficient reason for support, even if they

harbor concerns about the economy. This dynamic demonstrates how a political campaign can strategically shift the focus of public debate to its areas of strength to mitigate weaknesses.

The "Chowkidar" Campaign

The "Chowkidar" campaign serves as a fascinating case study in modern digital politics. The campaign was initiated by Modi and his supporters to position him as a "watchman" against corruption.¹ However, the data shows that this campaign was immediately and effectively subverted by the opposition.

- **Rapid Subversion of the Narrative:** The phrase "chowkidar chor" (watchman is a thief) became a popular counter-narrative, turning the incumbent's symbol of strength into a point of ridicule.¹ This rapid re-framing of the term highlights the fragility of digital political messaging. A symbol intended to project an image of honesty and guardianship was quickly re-appropriated to suggest the opposite.
- **The Digital Battle for Symbols:** This dynamic underscores that digital political campaigns are not static broadcasts but dynamic, conversational battles over symbols and meaning. A successful campaign must anticipate and prepare for counter-narratives from the outset. The reception of the "Chowkidar" campaign serves as a powerful example of how quickly and effectively a narrative can be turned against its creator in a highly polarized digital environment.

6. Visualizations and Tabular Data

The following tables and figures provide a quantitative summary of the qualitative observations detailed above.

Table 1: Overall Sentiment Distribution

Category	Count	Percentage
----------	-------	------------

Positive (1)	167	45.4%
Neutral (0)	114	31.0%
Negative (-1)	87	23.6%
Total	368	100%

Note: Data derived from the category column of X data.csv.¹

The data in Table 1 confirms the highly polarized nature of the discourse. The combined percentage of positive and negative posts (79%) is significantly higher than the neutral category (21%), indicating that most online conversations are not neutral but are driven by strong expressions of support or opposition.

Table 2: Top Keywords by Sentiment Category

Positive Keywords	Negative Keywords	Neutral Keywords
modi	modi	modi
india	modi govt	vote
bjp	congress	campaign
sir	modi and shah	bjp
new india	rahul	kar
proud	modi government	anil kapoor
great	corruption	supports
vote for modi	hate	appeal

leadership	liar	election
development	anti	appeals

Note: Data derived from a frequency analysis of the clean_text column, focusing on recurring themes and proper nouns.¹

Table 2 clearly illustrates the thematic differences in sentiment. Positive posts predominantly use terms like "proud," "great," and "leadership," creating a narrative of national pride and strong governance. Conversely, negative posts are saturated with critical terms like "corruption," "hate," "liar," and "anti," and frequently tie the incumbent to specific issues or accusations. This thematic separation reinforces the observation that different types of language and issues drive different emotional responses.

7. Conclusion and Strategic Recommendations

The sentiment analysis of the provided data points to a political landscape that is profoundly polarized and emotionally charged. The public discourse on social media is less about reasoned policy debate and more about the affirmation of a political identity, centered largely around the figure of Narendra Modi. The incumbent's success in leveraging nationalistic and strong leadership narratives has proven effective in mitigating criticisms related to tangible economic and social issues. The opposition, conversely, has struggled to move beyond a reactive, anti-Modi stance, which has made it difficult for them to establish a compelling and independent brand.

Based on this analysis, the following recommendations can be made:

For the Incumbent's Strategist:

- Narrative Control:** Continue to frame national achievements, be they military or scientific, as a testament to strong leadership and national pride. The success of this strategy is evident in the data, where events like the anti-satellite test or the Balakot air strike are consistently linked to Modi's personal political will.¹ This emotional appeal is a potent defense against policy-level criticism.
- Proactive Messaging:** While the existing strategy is strong, the rapid co-option of campaigns like "Chowkidar" suggests a need for a more proactive approach to

counter-narratives. Future campaigns should be developed with an explicit understanding of how they can be subverted and with a pre-planned strategy to neutralize such attempts.

For the Opposition's Strategist:

- **Build an Independent Brand:** The analysis highlights a critical weakness in the opposition's reliance on an anti-Modi narrative. The discourse is plagued by questions like "not modi then who?"¹, which shows a leadership vacuum. The opposition must build a strong, unifying brand around a leader and a clear, positive vision that is not solely defined by opposition to the incumbent.
- **Focus on Tangible Issues:** The negative sentiment in the data is strongly linked to economic issues such as joblessness and rising prices. A more effective strategy would be to focus relentlessly on these tangible policy failures, offering credible and well-defined solutions that resonate with the public's day-to-day concerns. The Nyay scheme, for instance, is a good example of a policy that attempts to address these issues directly and should be a central focus.¹

This report provides a snapshot of a complex and dynamic political environment. A future, more comprehensive analysis of a time-series dataset would be invaluable for understanding the causal relationships between specific events and shifts in public sentiment, providing a more robust foundation for long-term strategic planning.

Works cited

1. X data.csv