

Chowkidar Campaign: What is the Twitter Sentiment?

Hemanth Bharatha Chakravarthy

2019-03-21

Analysis of Twitter Reactions to the BJP's #Chowkidar Campaign

I analyzed a mixed sample of the most popular and most recent **one hundred and eight thousand tweets** in English from the last three days (as of 20 March, 2019, 10:07 pm EST) on the Bharatiya Janata Party's #MainBhiChowkidar campaign.

Sentiment Analysis: Were the Tweets Positive or Negative?

Comparing the number of positive tweets against the number of negative tweets.

Unsurprisingly, there are more positive tweets than negative. Surprisingly, the difference between positive and negative tweets is a small 4.93%. Now, this could possibly be explained by this analysis only containing English tweets and that Twitter users also tend to be more urban, educated, younger, and relatively elite. But, while I do expect a larger tendency towards positivity, I believe that this small window of difference does confirm our expectations of at least a somewhat closely contested and messy election.

How did Sentiments Change with Time in the Last Three Days?

As I expected, sentiments have been pretty stable though they seem to have been more positive on the 19th. What's interesting about this plot is the increased negativity and backlash against the campaign starting around 6 am on the 20th in American time. This is interesting because this roughly coincides with when the news about Nirav Modi's arrest in London was released. As you can see from the above wordcloud, words like Nirav Modi and arrest immediately become some of the most tweeted words.

Now, I expected a generally positive reaction to Nirav Modi's arrest. However, the data shows that Nirav Modi's arrest actually caused an increase in negative tweets about the #Chowkidar aka #MainBhiChowkidar campaign. A lot of tweeters seemed to believe that this is an election stunt and suspected the convenient timing of the arrest for the BJP.

A Final Developer's Note

Unfortunately, there seem to be no existing machine learning based APIs or CRAN packages to deal with Hindi Tweets, so I'm ignoring them. I considered translating and then analyzing, but this seems to have too broad a confidence interval and Google Translate API is too expensive for me. This obviously makes this analysis biased to some extent. Furthermore, even for English sentiment analysis, there is a non-zero margin of error. However, given my rather large sample size, this margin should be adjusted for. My analysis also does not account for paid tweets. So, the positive skew might be caused in some part by the BJP tech cell's tweeting. However, given that I scraped a mixed sample of popular and recent tweets and given that paid tweets are unlikely to be the most popular ones, this skew should be mitigated.

You can look at the code (in R) in my GitHub (@b-hemanth) under my repository `my_page`.

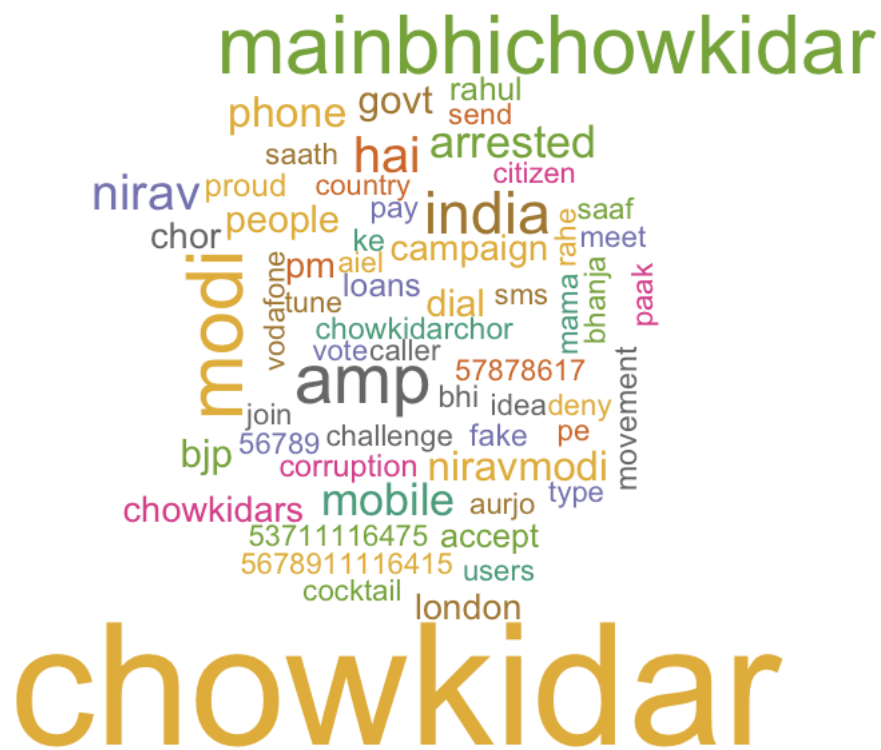


Figure 1: Word Cloud

Sentiment Analysis of Tweets About the #Chowkidar Campaign

Analyzing a Mixed Sample of 108,000 of the Most Popular and Most Recent Tweets

Sentiment	Percentage
Positive	54.93%
Negative	45.07%
Total	100.00%

Data from Twitter

Figure 2: Sentiments Table

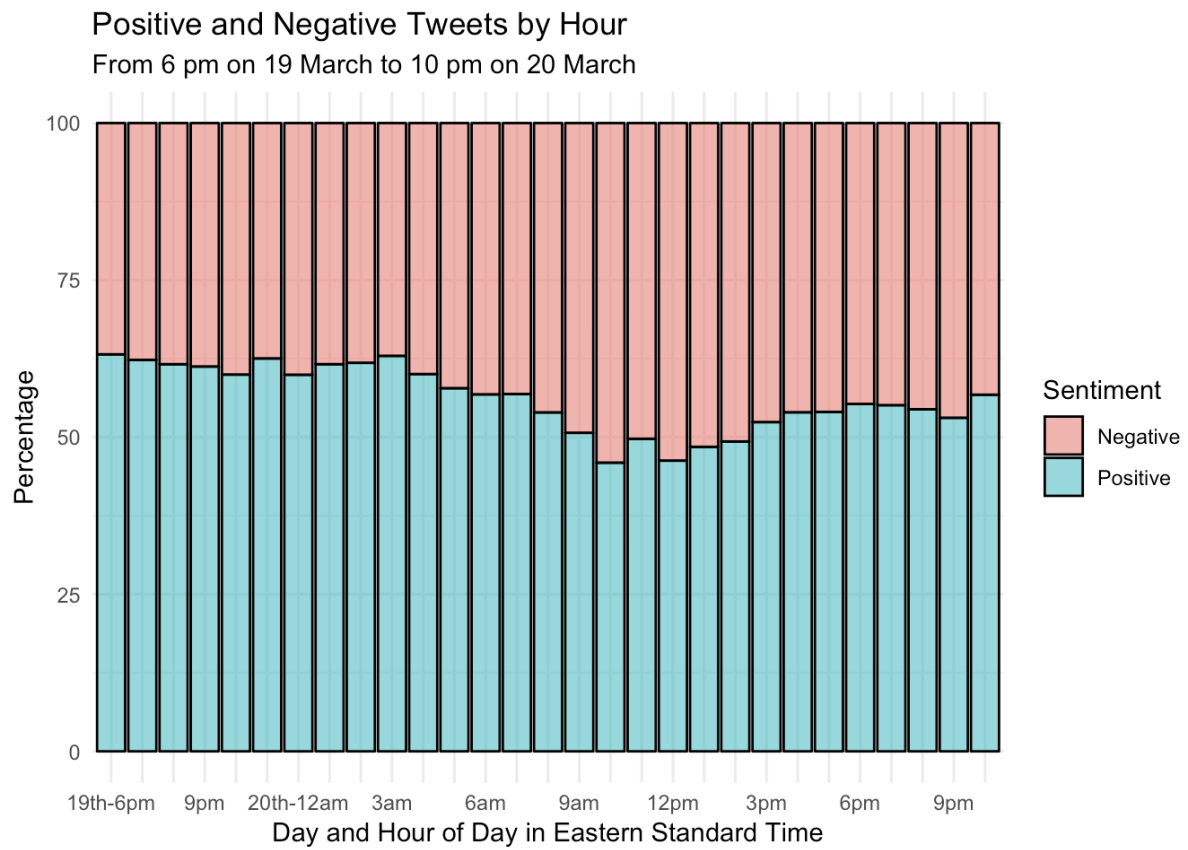


Figure 3: Sentiments Across the Day