

# MATH513 Big Data and Social Network Visualization

10883408

2023-12-05

# Introduction:COVID-19 Discussions on Twitter in India (February-March 2021): A Detailed Analysis

## Objective

- ▶ To get useful information by using a complete method that includes data visualization techniques and sentiment analysis.

## Approach

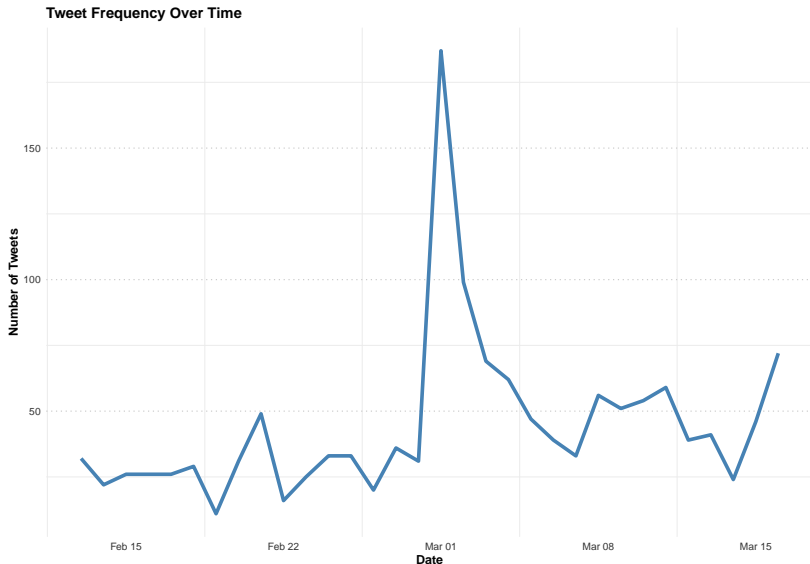
- ▶ Focus on using a variety of data visualization methods to show patterns, trends, and important measures in the Twitter dataset.
- ▶ Apply sentiment analysis to identify the emotional tones of the insights .

## Analytical focus

- ▶ Analysis of the frequency and intensity of conversations.
- ▶ Identification of prominent themes within the dataset.
- ▶ Evaluation of changes in sentiment during the provided timeframe.

# Tweet frequency over time

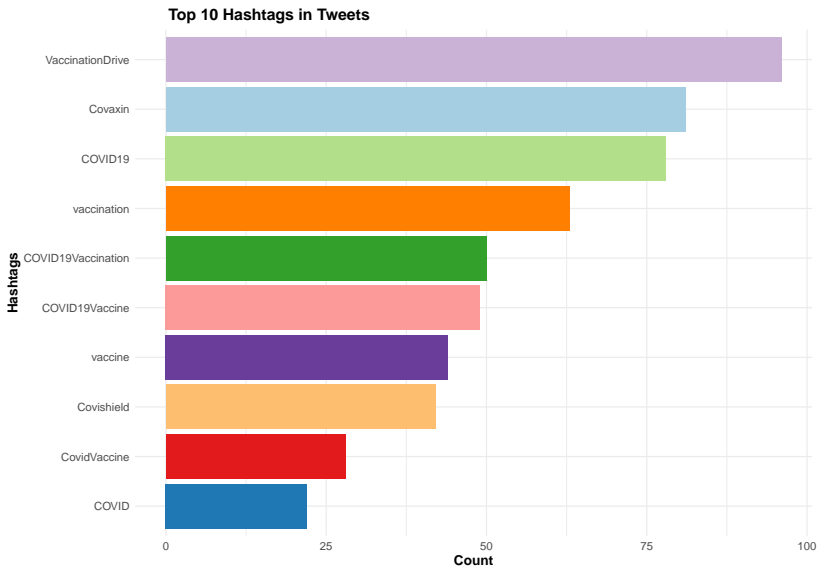
- ▶ The graph spans from February 13, 2021, to March 16, 2021.
- ▶ Huge increase in the volume of tweets on March 1, 2021.



# Why maintain , English only tweets

- ▶ Language Proficiency
- ▶ Resource Constraints
- ▶ Consistency and Comparability
- ▶ Availability of Tools and Libraries
- ▶ Research Focus

# Top hastags in Tweets

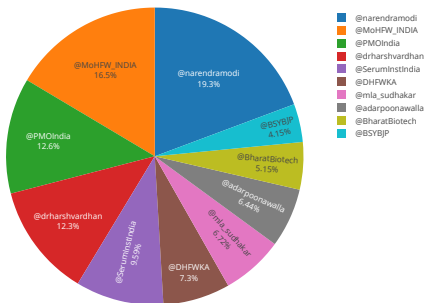


# Top Mentioned people within the period

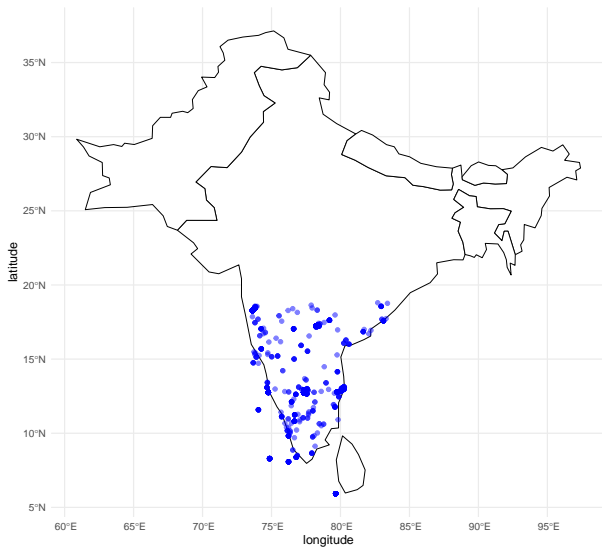
Mostly tagged people and Locations

- ▶ Narendra Modi
- ▶ MoH India
- ▶ Prime Minister office India

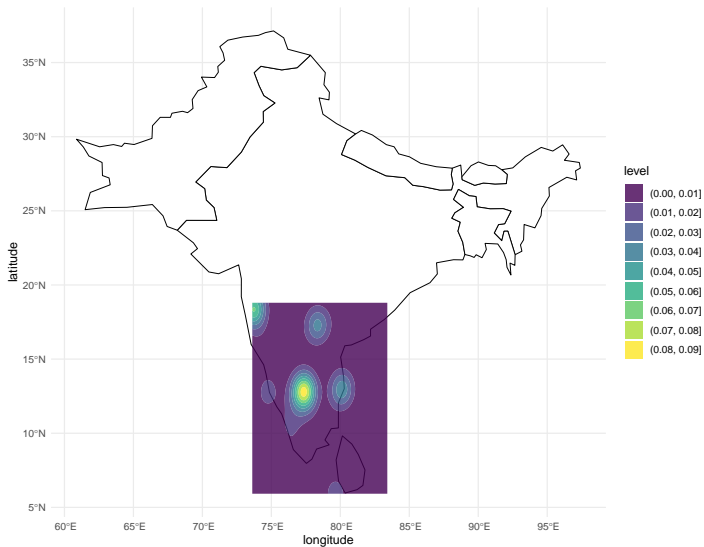
## PhantomJS not found. You can install it with `webshot::install`



# Geolocation Distribution

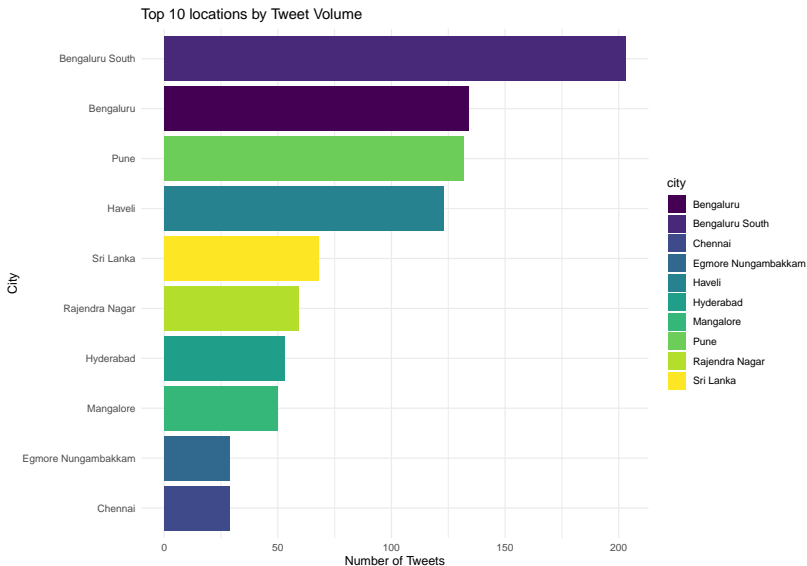


# Heatmap

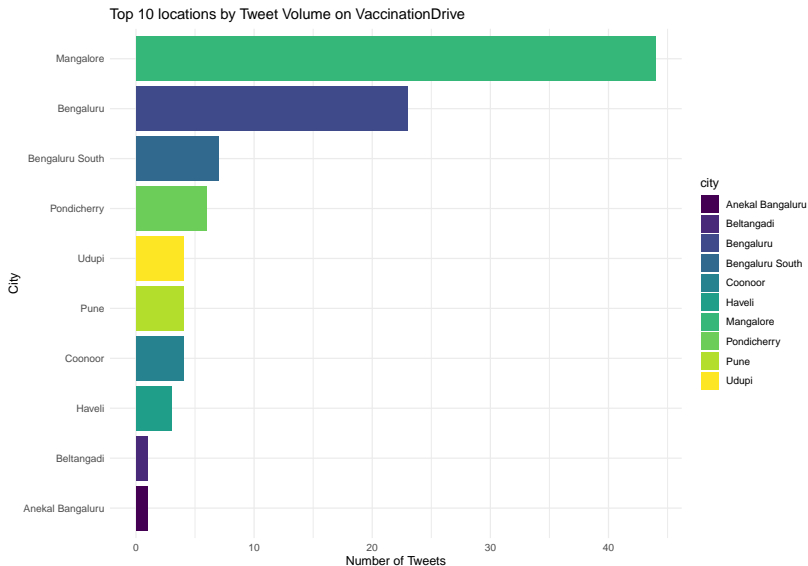




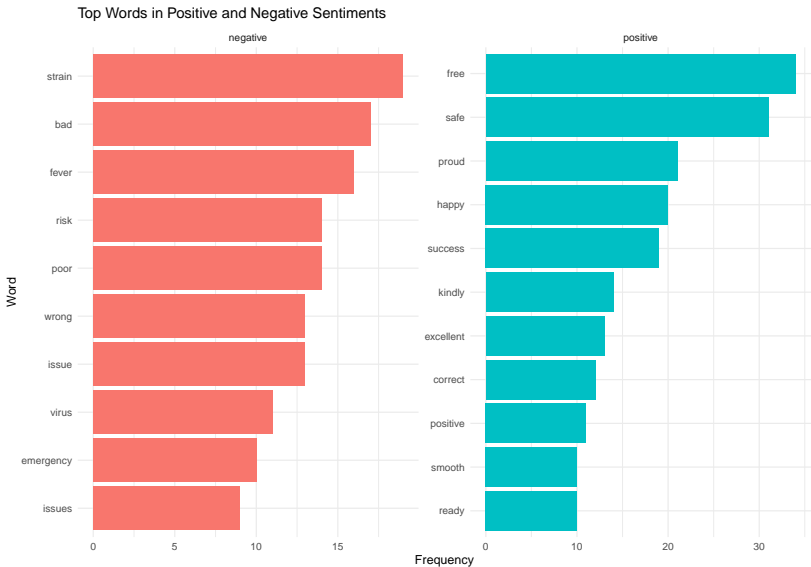
# Locations by tweet volume



# Locations by tweet volume and Vaccination Drive



# Sentiment Analysis



# The T-Test Applicability

- ▶ Assumption of Normality- The t-test assumes normally distributed group data. This is especially important for small samples.
- ▶ Sample Size- T-tests work with small and large samples, but smaller samples are less reliable.
- ▶ Outliers- When the sample size is small, outliers can have a big effect on the results of a t-test.

## Overall Conclusion

- ▶ Diverse Conversations - reflecting public opinion during the period.
- ▶ Trends and Patterns - helps with a visual understanding of ongoing discussions.
- ▶ Emotional Tone - showed a mix of positive and negative sentiments.
- ▶ Analytical Focus - conversation frequency, intensity, and thematic analysis helped clarify how people communicated.
- ▶ Geographical Distribution - highlighted South Asian discussion patterns.
- ▶ Qualitative Perspective - highlighted user emotions and attitudes over the specified timeframe.
- ▶ Stakeholder Insights - gives stakeholders a complete picture of Indian Twitter users' COVID-19 discussions.

# Suggestions

- ▶ Comparative Analysis - Compare the analyzed time frame to other relevant periods or countries to find unique patterns or trends in India.
- ▶ Public Health Impact - Work with public health experts to link online discussions to indicators.
- ▶ Multimodal Analysis - Add tweet images and videos to the analysis to better understand how multimedia affects discussions.