

# Analysis of Live Stream and Collected Tweets for U.S State Primary Related Data using R

## [1] Data Collection and Streaming

The data was collected for analyses using the streamR R library, with the search terms pertaining to the US Primary elections and popular presidential candidates. The tweets were collected into a *json* format and then analyzed for mentions of the most popular candidates.

The data was analyzed for the most popular content (by the number of retweets) and then the locations of these tweets taken and plotted on to a global map.

The analysis was done twice, once for the entire data for the data collected over a week, and once for the day analysis.

For streaming of live twitter data, a new stream connection is opened every 5 seconds and new tweets are collected. Once the tweets have been parsed, they are then analyzed to find new mentions or the locations of the new tweets generated.

## [2] Shiny UI Structure

The UI was structured as multi tab with user able to choose between mention analysis or location analysis. The user could also choose between Day Analysis, Weekly Analysis or Analysis of the live streaming tweets.



The search parameters can be changed to make the search more exhaustive but this would make the application slower and unresponsive.

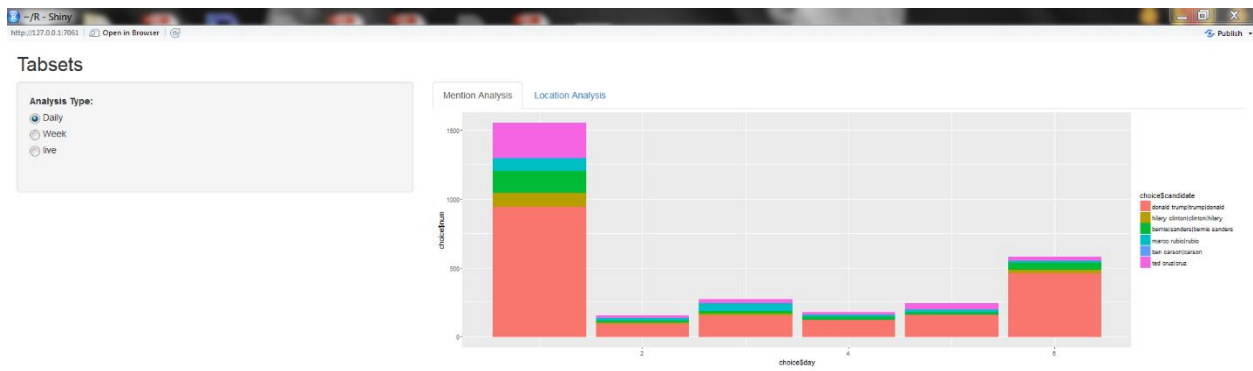


Fig: Mentions for popular candidates. (Per Day)



Fig: Mention Analyses (Live Stream)

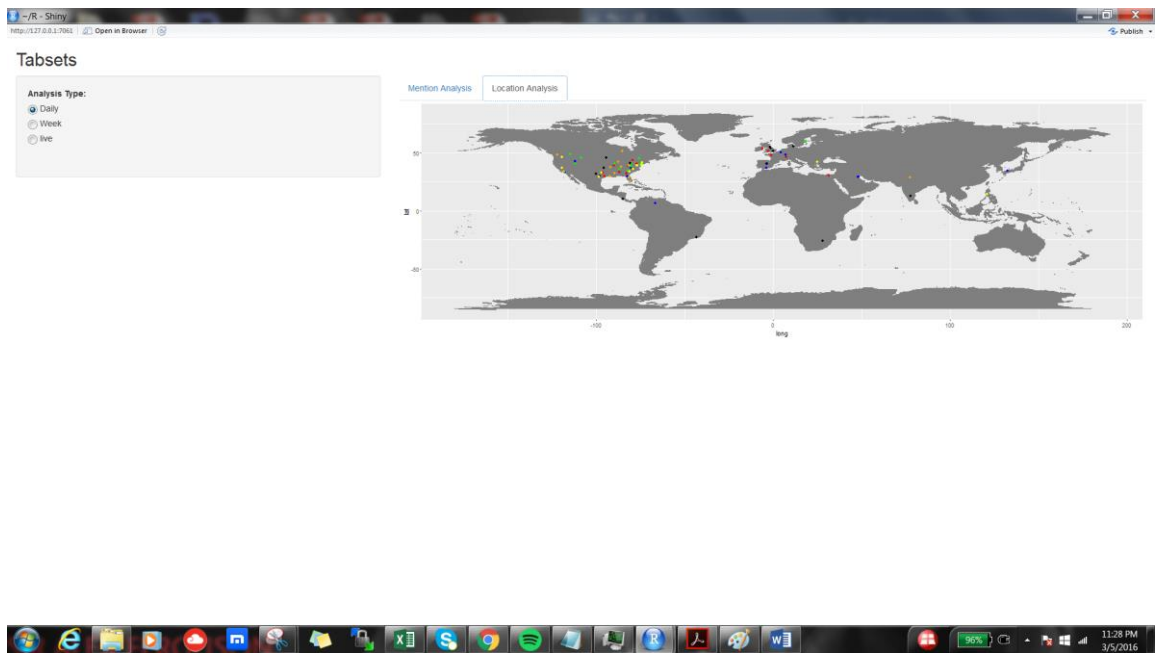


Fig: Location Analysis (Day Data)

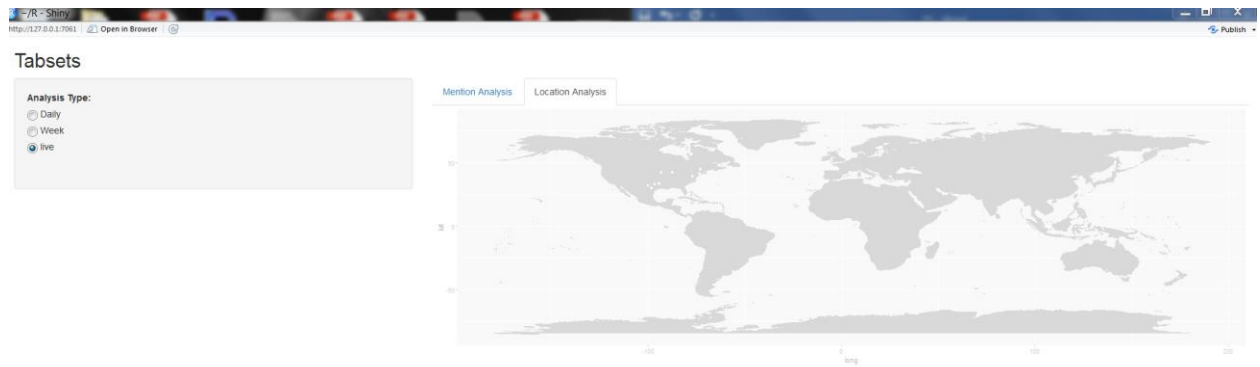


Fig: Location Analysis (Live Stream)