

CSE 587: Project 1

Problem 5 - Stream Processing

Lalith Vikram Natarajan, UBIT ID: lalithvi, Person No: 50169243

March 5, 2016

1 Objective

Create UI to access results

2 Data Source

Real Time Streaming Twitter Data was the dataset used.

3 UI Design

R Shiny was used to design the interface of the streaming application. The three main parts of the UI are the Map, a div containing relevant information about the dataset and a text input. The map is generated using the leaflet package in R. The tweets are streamed using the streamR package.

4 Analysis

To display the data once the user has logged in, I had already collected certain tweets which are then used to populate the map. The tweets are about Oscars. You can find the screenshot of the application below.

Twistreamer



On clicking the circle marker in the application, the user is able to see the tweets used to generate the location. This is shown below

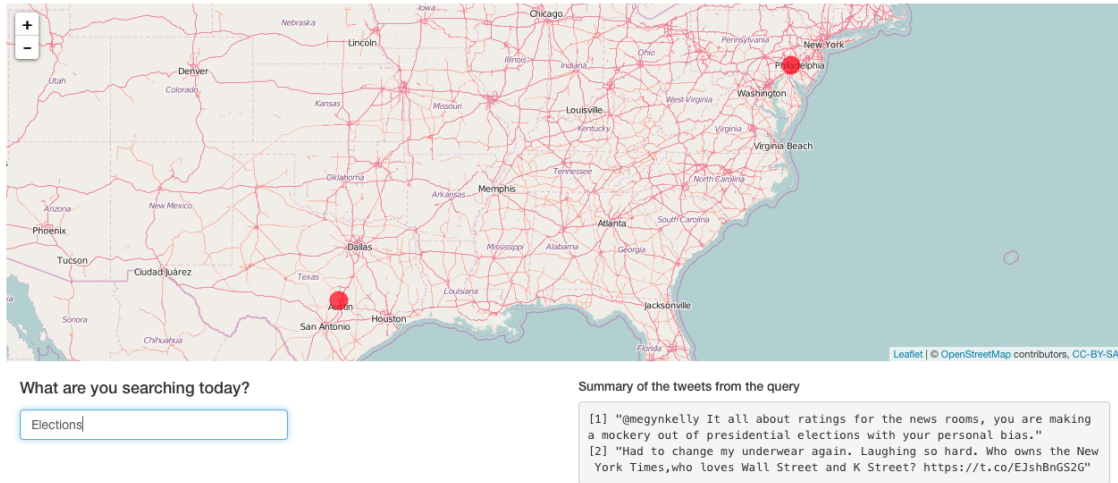
Twistreamer



After the user enters a value in the text input, the tweets are streamed for 120 seconds and

the results displayed. Note that the information div also changes pertaining to the tweets.

Twistreamer



With enough location data, the user will be able to figure out the spread of the phenomenon, for example while gathering data, I obtained 448000 tweets within ten minutes for the term Oscars on 28th of February. By spreading them over the map, we would be able to figure out what effect the particular incident had on that day and the locations in which it is popular. This being streaming data, would be very useful for figuring out which areas have the most influence (when running an advertising campaign, for example)

5 Drawbacks and Future Developments

The main drawback of this application is the non-availability of tweets with proper coordinates. This affects the overall number of tweets. When the streaming happens for a longer time, the user will get impatient. Also the information div could include the summary of tweets or a analysis, but the non-availability of tweets affected that decision as well. This should be the future directions in which the application should progress.