Tweetstorm Operation

? THE PROBLEM

The problem that has been addressed by this piece of work is to develop understanding of how influential social media can be on world events. To test our platform we have used the 2016 Presidential Election for the model.

> THE APPROACH

(1) Over 10 million tweets were collected using Twitter ID's related to key events during the 2016 US Presidential election, these ID's were used to extract the full tweets from the Twitter API. (2) Tweets were run against our sentiment and polarity model. (3) Once the full dataset was created queries were generated in order to develop datasets for visualization. (4) Visualizations were then developed including geographic mapping, polarity and subjectivity over time and emoji analysis. (5) Finally a platform was developed to walk through the results of the work.

- Data Collected
- Models run
- Data queries generated
- Visualizations developed
- Final presentation developed

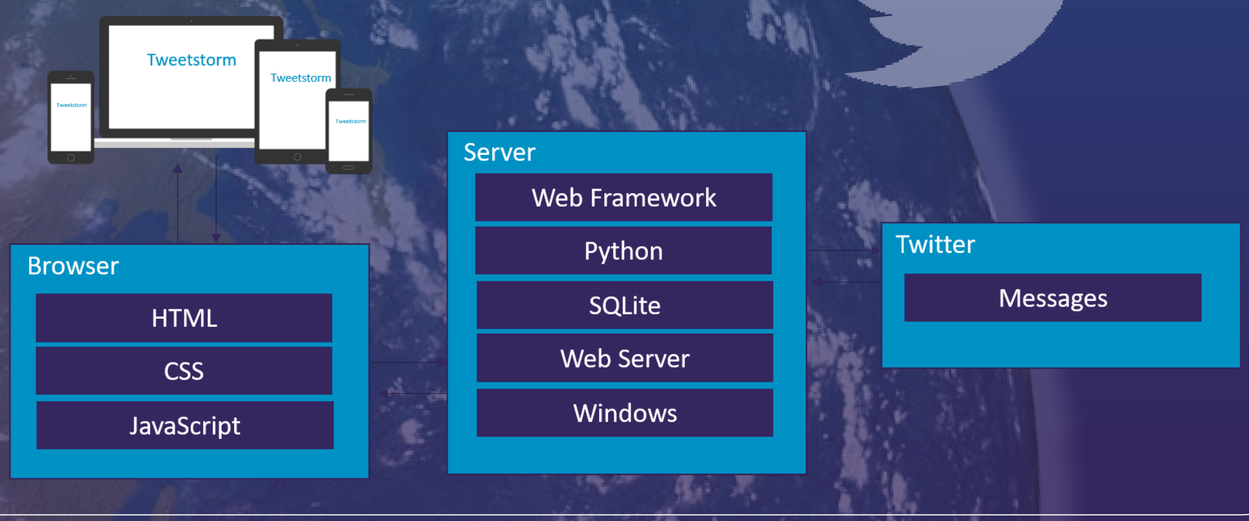
Our Data

Over 10 million tweets from stages of the 2016 US Presidential Election were obtained using the ID's captured by the Harvard Dataverse (Littman, Wrubel, Kerchner, 2016) and then calling the Twitter API to hydrate the IDs to obtain the full tweets. They were then run through a series of cleansing processes which were conducted to remove tags, links, and non-useful labels including "RT" for retweet.

! Experiments and Results

The Experiments conducted covered the areas of classification for the polarity and sentiment models including the classifying of emoji's: the key learning from this was the need for cleaning of the data for better model prediction. Testing of the visualization: The key learning was identified as a need to add filter features to these in order to make better sense of them including within the force directed graph.

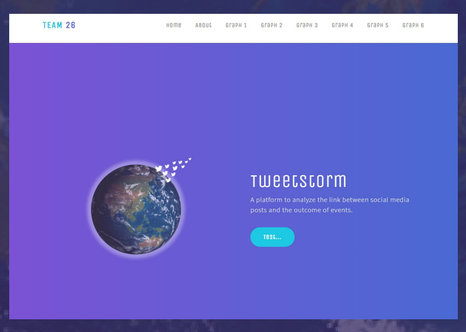
{ Technology Diagram



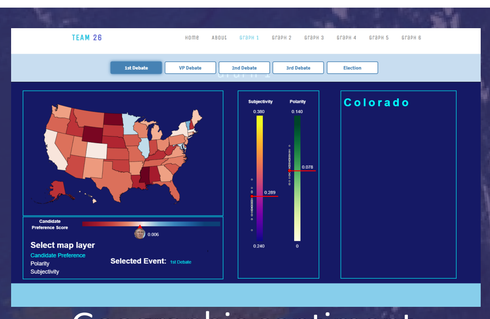
} Evaluation

Right from the outset the Twitter sentiment analysis showed that Donald Trump was the more popular candidate. Whilst polling may not have suggested this, except for one point through the election cycle, it appeared that Donald Trump was the most likely candidate using this model.

+ Showcase

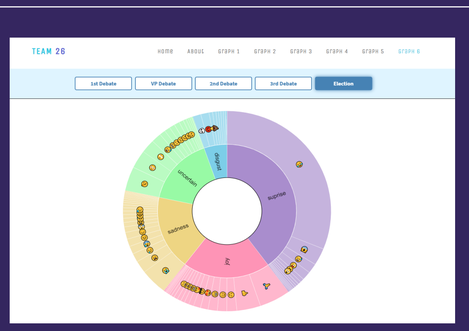


Hero Image

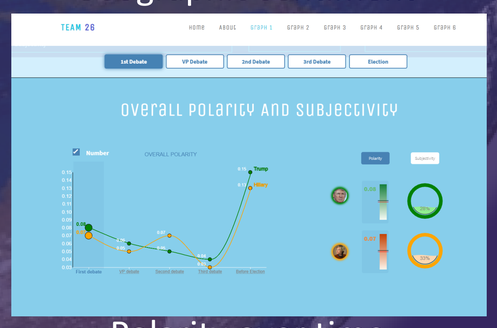


Geographic sentiment

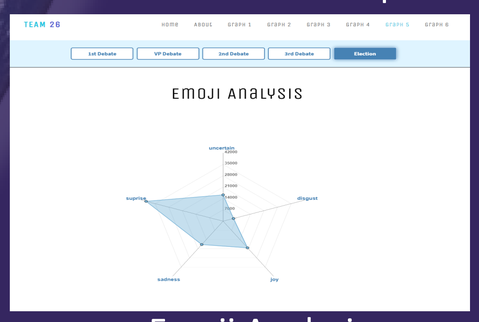
Force Directed Graph



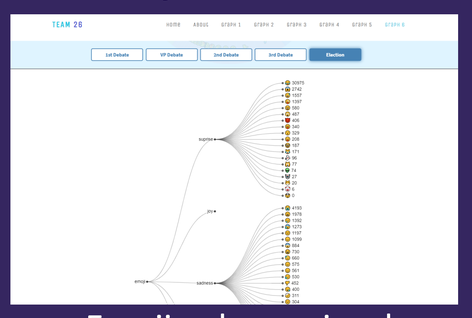
Emoji break-down



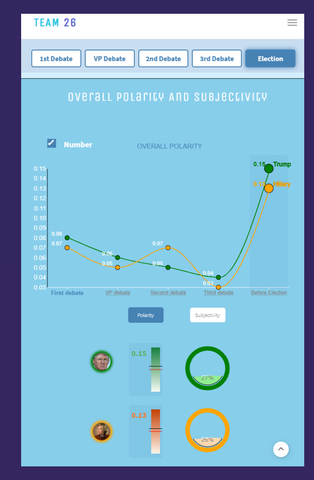
Polarity over time



Emoji Analysis



Emoji volume visual



Mobile Responsive View

Jeffrey Lee (jlee3703@gatech.edu) | Minhao Leong (mleong7@gatech.edu) | Jinchao Lin (jlin453@gatech.edu)
Jun Xiong Tan (jtan318@gatech.edu) | Dr Matthew Tinker (mtinker6@gatech.edu)